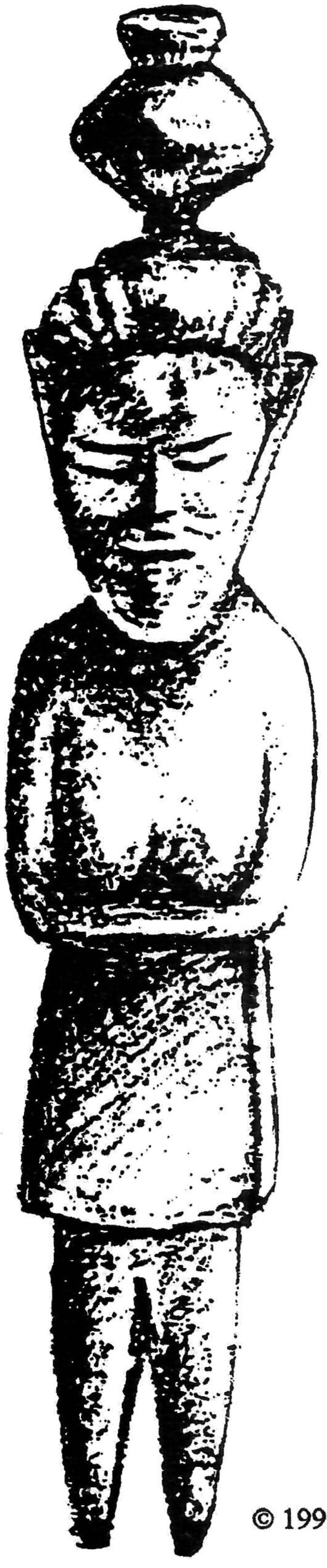
MOTHER TONGUE

JOURNAL OF THE ASSOCIATION FOR THE STUDY OF LANGUAGE IN PREHISTORY



Issue V, December 1999

CONTENTS

The Austric Languages

3 The Austric Denti-Alveolar Sibilants: *LaVaughn H. Hayes*
Comments by Discussants:
15 *Václav Blažek*
19 *Robert Blust*
23 *George van Driem*
29 *Harold C. Fleming*
33 Responses to Discussants: *LaVaughn H. Hayes*

The Basque Language: Part II

41 A Comparison of Basque and (North) Caucasian Basic Vocabulary: *John D. Bengtson*
Comments by Discussants:
59 *Václav Blažek*
63 *W. Wilfried Schuhmacher*
65 *Paul Sidwell*
71 *R.L. Trask*
87 *Edward J. Vajda*
93 *Paul Whitehouse*
99 Response to Discussants: *John D. Bengtson*

The Sumerian Language: Part II

109 An Enquiry about Sumerian Words: *Liny Asoka Srinivasan*
119 Comments on Srinivasan: *Michael Witzel*
141 More on Possible Linguistic Connections of The Sumerians: *Igor M. Diakonoff*
145 Response to Diakonoff: *John D. Bengtson*

Environmental Influences on Human Language

147 A Discussion of Climatic Influences on Language: *G.R. (Randy) Foote*
161 Bipeds, Tools and Speech: *Marc Verhaegen & Stephen Munro*

Prehistory of the Americas

169 The Evolution of Archaeological Perceptions of the First Americans: *Alvah M. Hicks*

Editorial

173 In Honor of Jürgen Pinnow: Comparanda Dene-Caucasica *John D. Bengtson*

© 1999 by the Association for the Study of Language in Prehistory
ISSN: 1087-0326 *Mother Tongue* (Journal)

MOTHER TONGUE Issue V, December 1999

OFFICERS OF ASLIP

(As of April 15, 2000. Address appropriate correspondence to each.)

President:

Michael Witzel witzel@fas.harvard.edu
Dept. of Sanskrit and Indian Studies
Harvard University
2 Divinity Avenue
Cambridge, MA 02138
U.S.A.

Vice-Presidents:

John D. Bengtson john.bengtson@co.hennepin.mn.us
156 15th Avenue NE
Minneapolis, MN 55413
U.S.A.

Roger W. Wescott whilro@aol.com
16-A Heritage Crest
Southbury, CT 06488
U.S.A.

Secretary-Treasurer:

Peter Norquest Norquesp@U.Arizona.edu
1632 Santa Rita Avenue
Tucson, AZ 85719
U.S.A.

Long Ranger Newsletter Assistant Editors: G.R. Foote, Los Angeles, CA
Alvah M. Hicks, Creston CA

Technical Advisor: Brita M. Bengtson, Minneapolis, MN

BOARD OF DIRECTORS

Ofer Bar-Yosef (Peabody Museum, Harvard)
Allan R. Bomhard (Charleston, SC)
Gyula Décsy (U. of Indiana)
Frederick Gamst (U. of Massachusetts)
Phillip Lieberman (Brown University)
Anne Windsor Beaman (Brookline, MA)
Ronald Christensen (Lincoln, MA)
Harold C. Fleming (Boston University)
Mary Ellen Lepionka (Gloucester, MA)
Jan Vansina (Madison, WI)

COUNCIL OF FELLOWS

Raimo Anttila (U.C.L.A.)
†Igor M. Diakonoff (St. Petersburg)
Ben Ohiomamhe Elugbe (University of Ibadan)
Dell Hymes (University of Virginia)
Winfred P. Lehmann (University of Texas)
†Karl-Heinrich Menges (University of Vienna)
Vitaly Shevoroshkin (University of Michigan)
Luigi Luca Cavalli-Sforza (Stanford University)
Aharon Dolgopolsky (University of Haifa)
Joseph H. Greenberg (Stanford University)
Sydney Lamb (Rice University)
Daniel McCall (Boston University)
Colin Renfrew (Cambridge University, UK)
Sergei A. Starostin (Russian State University)

(† Deceased members, pending special election)

Introduction to *Mother Tongue* V

By Roger W. Wescott

The disagreements among the twelve authors of the first nineteen selections in MT-V seem to me to spring primarily from the philosophical divergence between absolutists and relativists. The absolutists appear to regard some genetic connections between languages as indisputable and others as inconceivable. The relativists, by contrast, tend to regard all such connections as possible, but only some as probable. The relativists, moreover, seem to treat probable affiliations as differing in degree, some being more probable than others.

Believing that all relations are, by definition, relative, I belong, predictably, in the ranks of the relativists. And, as a monogenist, I further consider all languages to be genetically related. For me, the question to be addressed regarding most controversial linguistic genealogies is one of subgrouping: which languages are closely related, which distantly, and how closely or distantly? Consequently, I treat no linguistic evidence as either conclusive or beyond consideration.

The question of scientific evidence for postulated cognation depends, clearly, on the definition of science. For mathematicians, science consists primarily of logic. For chemists, science consists primarily of experiment. For linguists, science is rarely either of these. But what is scientific linguistics? I hope that a future issue of *Mother Tongue* may be devoted, at least in part, to a discussion of this question.

The only section of MT-V in which polemics are not conspicuous is the last one, concerning environmental influences on language. Yet Verhaegen and Munro may take exception to my having characterized their view of early human evolution as "aquaticism." The term that Verhaegen has more often employed in recent writings is "aquarborealism." Aquarborealism is the hypothesis that, when Pliocene hominoids began venturing beyond tropical forests, they did most of their foraging in shallow waters but sought rest and refuge in the trees, avoiding grasslands until they had developed substantial tool-using and weapon-wielding capacities – presumably in the Pleistocene Epoch.

The Austric Denti-alveolar Sibilants

La Vaughn H. Hayes

1. Introduction.

1.1. Discovering proof of linguistic genetic relationship is one of the most challenging tasks of the historical linguist, and the holy grail of this enterprise is irrefutable proof of any sort. Finding genetic proof is seldom easy, and the more distant the relationship, the more difficult the task becomes. On the one hand, the probative evidence required is complex and multi-dimensional, with lexical, phonological, and morphological elements needed, as explained in Austric I (Hayes 1992:147-9). On the other, what appears to be solid proof may be all too easily explained away as non-genetic change, such as borrowing, diffusion, or areal convergence.

1.2. As such things go, the proposal that Austroasiatic (AA) and Austronesian (AN) are genetically related has been afflicted by more than its fair share of probatory difficulties.[1] When Wilhelm Schmidt introduced the Austric hypothesis in 1906, he presented phonological and morphological evidence which most observers have found convincing. His lexical evidence has not been viewed in equally approbative terms, and for that reason, Austric still has not been generally accepted as a proven language family.[2]

1.3. In the writer's Austric series (Hayes 1992, 1997b, 1999), an effort has been made to present lexical evidence of sufficient quantity and quality to corroborate the Austric hypothesis, and in Austric III, the lexical gap may have been finally plugged by the presentation of 96 basic vocabulary correspondences.[3] Nevertheless, the writer is still keenly aware that this lexical proof, as promising as it may be, is subject to being explained away as the consequence of borrowing and the like. Fortunately, the writer has recently found, after nearly 20 years of looking, what may be the holy grail of the Austric hypothesis. Demonstration of this ostensibly irrefutable proof of Austric's existence is the subject of this presentation.

2. The Austric Denti-alveolar Sibilants and their Evolution.

2.1. Austroasiatic. AA */s/ and */z/ have undergone diverse changes, and this fact is evidenced in the diverse phonological correspondence seen in the Austric comparative glossary presented at the end of this article. In Hayes 1997a and 1997b, it was shown that their primary reflexes were retained as */s, z/, respectively, down to the MK and Munda subfamily levels, but subsequently merged as */s/. This */s/ later shifted to /h/ and often then to /0/ in many languages. It was also shown that the sibilants have undergone a number of environmentally conditioned changes. At the PAA level, they coalesced with other phonemes as palatal stops and sibilants and shifted to the palatal sibilants */ś, ź/, respectively, when contiguous to */i/. More recently, the AA sibilants or their reflexes have undergone other changes, some of which are shown in section 3.

---

[1] Abbreviations used here are AA (Austroasiatic), AN (Austronesian), MK (Mon-Khmer), MP (Malayo-Polynesian), MUK (Mường Khến), NK (Nyah Kur), OM (Old Mon), P (Proto-), PM (Proto-Mon), PMN (Proto-Mnong), PNB (Proto-North Bahnaric), PSB (Proto-South Bahnaric), PW (Proto-Waic), and VN (Vietnamese).

[2] As used here, Austric means only the language phylum comprising the AA and AN language families. Others have begun using Austric to mean a taxonomic unit comprising those and other language families; I do not agree with their position, but I am not familiar with their reasons for using Austric in that extended sense.

[3] 96 AA/AN correspondences. An additional 70 MK/AN or Munda/AN basic vocabulary correspondences were also presented.

Here, it will be shown that */z/ also changed at the PAA level to a denti-alveolar stop in the nasal cluster, i.e. */nz/ > */nd/ (*/ns/ > */nt/ may also have occurred, but no clear example has been found thus far). The reader will also detect in the glossary irregular AA correspondences, such as */z/:/j/, cf. PAA */zaqa/, Katu ʔəjɨɨh 'snag', which indicate still other changes to the denti-alveolar sibilants. These changes are not discussed here because it is believed that they took place at the Austric or an earlier level. They will be, however, the focus of another paper to be presented at a later date.

2.2. Austronesian. Otto Dempwolff (cf. Dahl 1973:101) did not reconstruct denti-alveolar sibilants for Proto-Austronesian, but the Formosan evidence, which he did not have available, makes it clear that the denti-alveolars */s, z/ are to be reconstructed at the PAN level. Paul K. Benedict (1976:155) shows that Proto-Austro-Tai */s, z/ > Formosan */s, z/, Indonesian (MP) */h, D/, respectively, and reconstructs elsewhere */s, z/ for Proto-Austronesian. To my knowledge, mainstream Austronesianists have yet to correct this discrepancy, and still use Dempwolff's */h/ and */D/ where */s/ and */z/, respectively, should be reconstructed at the PAN level. That notational practice will be followed here in the citation of AN proto-forms.

2.3. Austric. Since both Proto-Austroasiatic and Proto-Austronesian possessed the denti-alveolar sibilants */s/ and */z/ and the respective AA and AN reflexes of these proto-phonemes correspond regularly and recurrently, */s/ and */z/ may also be reconstructed at the Austric level.

3. Exemplary Data.

3.1. General. Single-example comparisons are cited in this section in order to make the sibilant correspondence maximally clear; additional lexical examples are given in the glossary.

3.2. Reflexes of Austric */s/.

Austric	**Austroasiatic**	**Austronesian**
*s	*s	*s > *h

Austroasiatic	**PAA**	**Austronesian**
Khmer ṛas 'root'	*ɣasi	AN *biRah 'alocasia (species)'
VN thɤm 'be fragrant, smell good'	*sarʔom	AN *ha[r]um 'aroma, scent'
Katu saak 'corpse'	*sa[ʔ]ak,	AN *hawak 'body'

Semelai cəruus ‘claw (nail)’	*caḷus	AN *[s]ilu[h] ‘fingernail’
Bahnar kəsaayʔ ‘sprinkle’	*saqi	AN *bahaq ‘flood(ed)’
OM sok ‘hair’	*suk	AN *buhuk ‘hair’
Katu takɔh ‘old man or woman’	*tunqas	AN *tuqah ‘old’
Katu saal ‘pound rice’	*sa[ḷu]	AN *haluh ‘pestle, pounder’
Semelai gnɔs ‘heart’	*ganosi	AN *sa(ŋ)guh ‘pith, sago’
Souei səŋkaal ‘skin’	*sa(n)qaɣ	AN *haka[r] ‘root’
Khmer sruup ‘swallow, sip’	*suɣup	AN *hiRup ‘sip’
VN (*[h]waːś >) vay ‘ancestor’	*rawasi	AN *ñawah ‘spirit, soul’
PM *pruus ‘squirt’	*buɣasi	AN *buRah ‘spray, sprinkle’
PM *ʔbuh ‘boil’	*(m)bus	AN *sebuh ‘develop steam’
Pearic sro(:)t ‘undress oneself’	*sarut	AN *hurut ‘stroke’
PM *gnis ‘canine tooth’	*g[a]nis	AN *gigih ‘tooth’
PW *rɨs ‘turtle’	*(kən)[ḷ,r]us	AN *peñuh ‘turtle’
Pacoh ŋeaih ‘count’	*ŋkasi	AN *zaŋkah ‘unit of measure’
PW *pes ‘sweep’	*(tam)pis	AN *ta(m)pih ‘winnow’

3.3. Reflexes of Austric */z/.

*z	*z > *s	*z > *D
Katu sənɛɛt ‘belt for skirt’	*zanit	AN *genDit ‘belt, girdle’
Pacoh sɨk ‘butt, gore’	*zok	AN *[t]un[D]uk ‘to bow, bend down/over’
Katu (High) ʔəsɤʔ ‘leaf’	*zaqa	AN *[d,D]aqan ‘branch’
Sedang kəsah ‘shoulder’	*zaG	AN *DaDah ‘breast, chest’
Bonda ɔm[5] ‘hatch egg’	*zəm	AN *DemDem ‘brood, hatch’
Sengoi sʌr ‘shove, push aside’	*zoḷ	AN *sunDul ‘bump (into), push’
OM saŋ ‘conch’	*zaŋ	AN *qu(n)Daŋ ‘crustacea(n)’
PW *som ‘night’	*zəm	AN *DeDem ‘dark, black’
Ruc sak ‘dregs’	*zak	AN *[d,D]ak[ih] ‘dirt on skin’
Mon mih ‘body dirt’	*məz	AN *cemeD ‘dirty’
Pacoh kəsaaŋ ‘make a sound’	*zaŋ	AN *genDaŋ ‘drum’
Chrau suk pələy ‘village’	*zuk	AN *Duk ‘dwelling place, residence’
Katu səŋ ‘hear’	*zəŋiɣ	AN *DeŋeR ‘hear’
Bonda siʔ ‘fever, pain’	*ziq	AN *pe[d,D]iq ‘hurt, smart, sting’
VN (*zay >) tay ‘at, in’	*zay	AN *Dayah ‘inland/interior’
Bahnar səlaam ‘in between’	*zaḷəm	AN *Dalem ‘inside’
Gutob sukuʔ ‘ladle’	*zuk	AN *ci(n)Duk ‘ladle, scoop’
Stieng saw ‘see’	*zaw	AN *tin[D]aw ‘look at closely’
Mon sapaa ‘back (of hand, foot)’	*zapal	AN *DapaN ‘palm, sole’
Pearic khsal ‘wind’	*zal	AN *huDan ‘rain’

---

4. */s/ > /0/ occurs in Bonda and other South Munda languages, possibly via an intermediate shift to */h/.

Bahnar sut ‘wipe’	*zoc	AN *kaDus ‘rub, scratch’
Pacoh saat ‘wipe off, rub’	*zat	AN *DasDas ‘rub off’
Sora okij ‘a little more’	*zekiq	AN *Dikiq ‘small, little’
Chrau sɤɤr ‘take out, up’	*zeγ	AN *DiRih ‘stand (up)’
Khmer phsaar ‘join (two members) with gum, glue or the like’	*zər	AN *pizer ‘to stick’
Sora sʔiːŋ-ən ‘a house’	*zeŋ	AN *[d,D]iŋ[d,D]iŋ ‘wall’
Khasi sum ‘bathe’	*zom	AN *[d,D]anum ‘water/fresh water’
Brou sɔɔy ‘tail’	*uzay	AN *huDay ‘worm’

In a few cases where AN */d/ is found, the AA evidence indicates that */D/ should be reconstructed.

Sengoi dʌr ‘burn, flame up’	*nzaR	AN *damaγ ‘light, resin, torch’
Bonda sɔmɔr ‘flame to rise’	*zamaR	AN *damaγ ‘light, resin, torch’
Rengao həra ‘duck’	*zara	AN *da[r]ah ‘pigeon/dove’
PM *(ʔa)da ‘duck’	*nzara	AN *da[r]ah ‘pigeon/dove’
Bahnar kəsɔh ‘spit’	*zaq	AN *ludaq ‘spittle’

In these examples, AA /c/ corresponds to AN */D/. The AA palatal probably reflects coalescence of a prefix-initial cluster, as in */z > s > [k]s > c/.

VN cɤp ‘(of heaven) lighten, blink’,	*z[ə]p	AN *han[D]ap ‘flicker’
Pacoh chut ‘scrub’	*zoc	AN *kaDus ‘rub, scratch’
MUK cɐt ‘wash clothes’	*zat	AN *DasDas ‘rub off’
Pacoh cuur ‘sleepy or sad eyes’	*zoγ	AN *tiDuR ‘sleep’

The absence of an AA reflex of */z/ in the following example suggests that the voiced sibilant may have been used as a suffix in AN.

Santali bulu ‘thigh’	*buḷu	AN *luluD ‘shin’

The following examples reflect changes of */z/ in the nasal-cluster environment to a voiced denti-alveolar stop, implosive, or retroflex according to the language involved. This type of change occurred first at the PAA level where */nz/ > */nd/. The nasal later dropped out, sometimes first conditioning the shift of */d/ to /ʔd/ or /ḍ/, in certain languages. In the post-PAA era, new nasal clusters were created, and the same shifts were often repeated, with */nz/ > */nd/, etc.

*z	*z > *nz > d, ʔd, ḍ	*z > *D
Sora duj ‘bend’	*nzoki	AN *[t]un[D]uk ‘to bow, bend down/ over’
Kharia kuṇḍɑʔ ‘husk’	*nzak	AN *DeDak ‘bran’
Kharia ḍaʔ ‘waist’	*nzaG	AN *DaDah ‘breast, chest’
MUK ʔdɐŋ ‘sunshine’	*nzaŋ	AN *daDaŋ ‘bright, shine’

Sora dum-dum ‘hatch egg’	*nzəm	AN *DemDem ‘brood, hatch’
Bahnar ?dɤr ‘stamp on’	*nzoḷ	AN *sunDul ‘bump (into)’
Khasi kajiŋ dum ‘night’	*nzəm	AN *DeDem ‘dark, black’
Mon (k)?daŋ ‘be sluggish’	*nzəŋ	AN *[d,D]eŋen ‘dazed, deafened’
VN ?dam ‘have an intense desire for’	*nzam	AN *qiDam ‘desire’
OM [t]a?dak ‘filth’	*nzak	AN *[d,D]ak[ih] ‘dirt on skin’
Sora kədiŋ-ən ‘a drum’	*nzaŋ	AN *genDaŋ ‘drum’
Bonda ḍikh ‘to stay’	*nzuk	AN *Duk ‘dwelling place, residence’
Brao ?dɔwŋ ‘hear’	*nzəŋiɣ	AN *DeŋeR ‘hear’
PM *p()?dəy ‘inside’	*nzay	AN *Dayah ‘inland/interior’
PW *dok ‘tray’	*nzuk	AN *ci(n)Duk ‘ladle, scoop’
OM kandar ‘wife’	*nzar	AN *Darah ‘maiden, virgin’
Khmu’ sndɛh ‘dish, plate’	*nzeh	AN *kenD[ih] ‘pitcher, water jar’
Pearic dus skip ‘rub’	*nzoci	AN *kaDus ‘rub, scratch’
Juang doko ‘sit’	*nzuk	AN *DukDuk ‘sit’
Sora kədur ‘snore’	*nzoɣ	AN *tiDuR ‘sleep’
Kharia daw ‘big flat sickle’	*nzaw	AN *maŋ[d,D,j]aw ‘war sword’
Chrau ?ding ‘wall’	*nzeŋ	AN *[d,D]iŋ[d,D]iŋ ‘wall’
Gutob dum ‘drown’	*nzom	AN *[d,D]anum ‘water/fresh water’
Pacoh ndɔɔy ‘little finger/toe’	*unzay	AN *huDay ‘worm’

In the following examples, the nasal cluster */ns/ was recreated after */z/ was devoiced, whereafter the sibilant shifted to a voiceless denti-alveolar stop, implosive, or retroflex.

*z > *s > *ns > t, ?t, ṭ	*z	*z > *D
Sedang rəta ‘chest’	*zaG	AN *DaDah ‘breast, chest’
PW *ktɐm ‘egg’	*zəm	AN *DemDem ‘brood, hatch’
Katu tuŋ ‘deaf’	*zəŋ	AN *[d,D]eŋen ‘dazed, deafened’
Kharia kɔṭɛ? ‘be dirty’	*zak	AN *[d,D]ak[ih] ‘dirt on skin’
Kharia tɑñ ‘weave’	*zali	AN *DanDan ‘plait ropes’
Khmer ṭus ‘rub’	*zoci	AN *kaDus ‘rub, scratch’
Khmer ṭaaw ‘sword’	*zaw	AN *maŋ[d,D,j]aw ‘war sword’
PW *[ntiŋ] ‘wall’	*zeŋ	AN *[d,D]iŋ[d,D]iŋ ‘wall’

4. The Irrefutable Evidence for the Austric Hypothesis.

4.1. Historical Background.

4.1.1. The dating and location of Proto-Austric are issues yet to be resolved. Since the Austronesianists claim that Proto-Austronesian was spoken by 6,500 B.P. (cf. Blust 1993:6), Proto-Austric can be tentatively dated to prior to that time. The Austric homeland seems most likely to have been somewhere in southern or southeastern China. The Austric dialect or dialect group which became Proto-Austronesian was presumably situated either in southeastern China or on the nearby island of Formosa. Austronesianists have long proposed one of these areas as the birth place of Austronesian. The Austric dialect/dialect group which became Austroasiatic was presumably located somewhere to the west or southwest of Pre-Austronesian, whence it drifted southwards into Indo-China; hence, Proto-Austroasiatic most likely evolved in southern China.

4.1.2. Given those circumstances, speculative as they may be, it seems probable that the respective Austric dialects had already lost contact before they evolved into Austroasiatic and Austronesian. This is especially true if Proto-Austronesian evolved on Formosa, for we have no reason to believe that Austroasiatic was ever located on that island. Thus, Austroasiatic and Austronesian probably evolved in isolation from one another, and it is unlikely that they shared vocabulary or other linguistic features during their developmental stages.

4.1.3. After their appearance, Austroasiatic and Austronesian apparently remained out of contact for some length of time, perhaps hundreds, even thousands of years. Eventually, some speakers of the MP subgroup of Austronesian returned to the mainland, and contact between the two Austric families was re-established. When and where this took place is also unknown precisely, but the coastal area of southern China or Vietnam is a likely location. However, Proto-Austroasiatic was probably long gone by that time; thus, the contact was with Mon-Khmer. It is known that the MK and MP languages have been in contact in Vietnam for at least two millennia, but the Munda subgroup of Austroasiatic possesses no AN loanwords as far as we can tell, hence has apparently never been in contact with any subgroup of Austronesian.

4.2. Indications of the Sibilant Evolution.

4.2.1. Throughout the period of the evolution of the Austric dialects into Austronesian and Austroasiatic, it is clear that Austric */s/ and */z/ were retained in both languages. They continued to be retained in Austroasiatic until well after the PAA era. However, when Austronesian split into its two primary subgroups, the sibilants were retained in Proto-Formosan, but shifted to */h/ and */D/, respectively, in Proto-Malayo-Polynesian. This split presumably occurred on Formosa or perhaps when Pre-Malayo-Polynesian moved south to the Philippines, with the shift to */h, D/ apparently occurring at or around the same time.

4.2.2. As indicated above, we have no reason to think that Austroasiatic and the Formosan languages were ever in geographic contiguity, hence also no reason to believe that these language groups ever exchanged vocabulary or anything else. It is thus only with Malayo-Polynesian that Austroasiatic has had contact and thus the possibility to exchange loanwords, and this contact apparently did not take place for a great length of time after the emergence of Proto-Malayo-Polynesian when MP speakers returned from the islands to the mainland and then only with the Mon-Khmer subfamily of Austroasiatic. By that time, the shift of PAN */s, z/ to PMP */h, D/, respectively, was ancient history.

4.2.3. The preceding observations allow us to draw two important conclusions about the AA/AN comparison. First, no AN loanword in Austroasiatic can possess a sibilant reflex of the Austric denti-alveolar sibilants. If a loanword has a reflex of the Austric sibilants, it must be a reflex of the */h/ or */D/ which replaced the PAN and Austric sibilants in Proto-Malayo-Polynesian (to my knowledge, */h, D/ have no sibilant reflexes in Malayo-Polynesian). If this constraint is not met, the word is not an AN loanword or the sibilant is not a reflex of the Austric sibilants.

Second, correspondent AA/AN lexical forms which possess corresponding reflexes of the Austric denti-alveolar sibilants, i.e. the various reflexes of */s/ or */z/ shown in sections 2 and 3 on the AA side, */h/ or */D/ on the AN side, must reflect forms inherited from a common ancestor. Common possession of inherited vocabulary is prima facie evidence of genetic relationship; hence, Austroasiatic and Austronesian are genetically related. This evidence can also be regarded as irrefutable proof of that relationship, because the regular and recurrent correspondence of AA */s/ and */z/ to PMP */h/ and */D/, respectively, cannot be accounted for in any other way.

5. Conclusion. The Austric denti-alveolar sibilants have evolved in distinctly different ways in Austroasiatic and Austronesian. As a result, a unique correspondence set, AA */s, z/ to PMP */h, D/, respectively, exists, and may be used to distinguish inherited vocabulary from borrowed vocabulary in these language families. That correspondence can also be interpreted as irrefutable proof that Austroasiatic and Austronesian are genetically related and descendants of a common ancestor, Austric, as proposed by Schmidt in 1906.

References[5]

Banker, John and Elizabeth and Mơ. 1979. *Bahnar Dictionary, Plei Bong-Mang Yang Dialect.* Huntington Beach:SIL.

Barker, Milton and Muriel. 1976. *Mường-Vietnamese-English Dictionary.* Microfiche. Huntington Beach:SIL.

Benedict, Paul K. 1975. *Austro-Thai: Language and Culture, with a Glossary of Roots.* New Haven:HRAF Press.

Bhattacharya, Sudhibhushan. 1968. *A Bonda Dictionary.* Poona:Deccan College Post-Graduate and Research Institute.

Benjamin, Geoffrey. 1976a. "Austroasiatic subgroupings and prehistory in the Malay Peninsula." *Austroasiatic Studies, Part I*:37-128.

Blood, Henry F. 1968. *A Reconstruction of Proto-Mnong.* Grand Forks:SIL-University of North Dakota.

Blust, Robert. 1993. "Beyond the Austronesian Homeland: The Austric Hypothesis and Its Implications for Archaeology." Paper presented at the Nov 1993 American Philosophical Society conference in Philadelphia, PA.

Costello, Nancy A. 1971. *Katu Vocabulary.* Saigon:Bộ Giáo Dục.

Dahl, Otto Christian. 1973. *Proto-Austronesian.* Scandinavian Institute of Asian Studies Monograph Series No. 15. Lund:Studentlitteratur.

Diffloth, Gérard F. 1980. "The Wa languages." *Linguistics of the Tibeto-Burman Area 5.2.*

---

[5] Abbreviations used in this section: MKS (Mon-Khmer Studies), SIL (Summer Institute of Linguistics).

Diffloth, Gérard F. 1984. *The Dvaravati-Old Mon Language and Nyah Kur. Monic Language Studies 1*. Bangkok:Chulalongkorn University.

Ferlus, Michel. 1974. "Lexique souei-français." *Asie Sudest Et Monde Insulindien V(1)*:141-59.

Ferlus, Michel. 1979. "Lexique thavung-français." *Centre de Langues Asie Orientale 5*:71-94.

Gradin, Dwight and Thong. 1979. *Jeh Vocabulary*. Huntington Beach:SIL.

Gregerson, Kenneth and Marilyn. 1977. *Rengao Vocabulary*. SIL.

Haupers, Lorraine and Ralph. 1991. *Stieng-English Dictionary*. Manila:SIL.

Hayes, La Vaughn H. 1992. "On the track of Austric, part I: introduction." *MKS XXI*:143–77.

Hayes, La Vaughn H. 1997a. "Another look at final spirants in Mon-Khmer." *MKS XXVI*:51–64.

Hayes, La Vaughn H. 1997b. "On the track of Austric, part II:consonant mutation in early Austroasiatic." *MKS XXVII*:13–41.

Hayes, La Vaughn H. 1999. "On the track of Austric, part III:basic vocabulary comparison." *MKS XXVIII*:1-34.

Headley, Robert K. Jr. 1977. "A Pearic vocabulary." *MKS VI*:69–149.

Jacob, Judith M. 1974. *A Concise Cambodian-English Dictionary*. London:Oxford University Press.

Means, Nathalie and Paul B. Means. 1986. *Sengoi-English, English-Sengoi Dictionary*, ed. Gordon P. Means. Toronto:Joint Centre on Modern East Asia, University of Toronto-York University.

Miller, Carolyn and John. 1967. *English-Brou Language Familiarization Manual*. Saigon:SIL.

Nguyễn Đình Hoà. 1971. *Vietnamese-English Student Dictionary*. Carbondale:Southern Illinois University Press.

Nguyễn Phú Phong, Trần Trí Dõi and Michel Ferlus. 1988. *Lexique Vietnamien Ruc-Français, Parler d'une minorité ethnique des montagnes de Quang Binh, Vietnam,avec Notes ethnographiques et introduction linguistique*. Paris:Sudestasie.

Pinnow, Heinz-Jürgen. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden:Otto Harrassowitz.

Rabel, Lili. 1961. *Khasi, a Language of Assam*. Baton Rouge:Louisiana State University.

Ramamurti, G.V. 1938. *Sora-English Dictionary*. Madras:Superintendent, Gov't Press. Reprinted 1986, Delhi:Mittal Publications.

Shorto, Harry L. 1962. *A Dictionary of Modern Spoken Mon*. London:Oxford University Press.

Shorto, Harry L. 1971. *A Dictionary of the Mon Inscriptions from the Sixth to the Sixteenth Centuries*. London:Oxford University Press.

Smalley, William A. 1961. *Outline of Khmu' Structure*. New Haven:American Oriental Society.

Smith, Kenneth D. 1972. *A Phonological Reconstruction of Proto-North-Bahnaric (Language Data, Asian-Pacific Series, No. 2)*. Santa Ana:SIL.

Theraphan L. Thongkum. 1984. *Nyah Kur (Chao Bon)-Thai-English Dictionary, Monic Language Studies 2*. Bangkok:Chulalongkorn University.

Thomas, David D. and Thổ Sáng Lục. 1966. *Chrau Vocabulary*. Saigon:Bộ Giáo Dục. Reprinted 1977, Huntington Beach:SIL.

Watson, Richard and Saundra and Cubuat. 1979. *Pacoh Dictionary*. Huntington Beach:SIL.

Wurm, S.A. and B. Wilson. 1983. *English Finderlist of Reconstructions in Austronesian Languages (Post-Brandstetter), Pacific Linguistics Series C33*. Canberra:Australian National University.

Austric Comparative Glossary

Austroasiatic	PAA	Austronesian
Jeh riayh, Khmer ṛas, Semai rəʔis 'root'	*γeqasi, *γaqis, *[qi]γasi	AN *biRah 'alocasia (species)', Formosan *buγasi 'sweet potato'
VN[6] thɤm 'be fragrant, smell good', PW *sʔɤm 'rotting', Sora serum 'to smell'	*sarʔom[7], *ca(r)ʔom	AN *ha[r]um 'aroma, scent'
Katu sənɛɛt 'belt for skirt', Chrau chɛ nich 'belt', VN nit 'garter, to tie'	*[zit][8], *zanit	AN *genDit 'belt, girdle'
Katu saak 'corpse', ʔəcak 'body', Bahnar kiak 'ghost, corpse'	*sa[ʔ]ak	AN *hawak 'body'
Pacoh sɨk 'butt, gore', Katu pəjuk 'lower head to butt', Sora duj 'bend'	*(n)zok(i)	AN *[t]un[D]uk 'to bow, bend down/ over'
Kharia kuṇḍɑʔ 'husk', Jeh dak 'bran', Khmer kandak 'dust which falls off husked rice'	*nzak	AN *DeDak 'bran'
Katu (High) ʔəsɤʔ 'leaf', ʔəjɨɨh 'snag', Mon saka 'toothstick, twig or slip of wood for cleaning teeth'	*za(n)qa, *jaqa	AN *[d,D]aqan 'branch'
Sedang kəsah, Kotua kəsɤh 'shoulder', Sedang rəta 'chest', Kharia ḍaʔ 'waist'	*(n)zaG	AN *DaDah 'breast, chest'
MUK ʔdɐŋ 'sunshine', Brou moət məndaaŋ 'sun', VN nɐŋ 'be sunny, the sun'	*nzaŋ	AN *daDaŋ 'bright, shine'
Bonda ɔm, Sora dum-dum 'hatch egg', PW *ktɐm 'egg'	*(n)zəm	AN *DemDem 'brood, hatch'
Chrau[9] chɨɨl 'push', Sengoi sʌr 'shove, push aside', Bahnar ʔdɤr 'stamp on, push with feet'	*(n)zoḷ	AN *sunDul 'bump (into), butt, knock against, push'
OM saŋ 'conch', Rengao kəsɔŋ 'shrimp', Khmer khyaŋ 'shellfish'	*zaŋ, *ñjaŋ	AN *qu(n)Daŋ 'crustacea(n)'
PW *som, Khasi kajiŋ dum 'night', VN ʔdəm 'dark'	*(n)zəm	AN *DeDem 'dark, black'
Mon (k)ʔdaŋ 'be sluggish', ktow	*(n)zəŋ(gən)	AN *[d,D]eŋen 'dazed, deafened,

---

[6] */s(V)r > sr > ś/ > VN /th/ [tʼ] (orthographic th).

[7] Replaces */s(ə)rom/ cited in Hayes 1997b:22.

[8] The root */zit/, which replaces */(n)dit/ cited in Hayes 1997b:29, is not evidenced unless in Bonda gɔsi 'wear cloth (by men)' and Sora gu.-sɑ:j 'cover (one's body) with cloth'.

[9] Chrau */c(ə)s-/ > /ch-/ (orthographic chh-). This change is seen in a number of MK languages, with */ch-/ often subsequently shifting to /s/.

k?daŋ ‘deaf (ktow ‘ear’)’, Katu tuŋ ‘deaf’, VN ŋən ‘look dumb-founded’		dizzy, numb, stunned’
VN ?dam ‘have an intense desire for’, zəm ‘lustful, sexy, lewd’, Sora kəyeem ‘be fond of, long for (as food, music, women)’	*nzam, *ñjam	AN *qiDam ‘desire/desiderative marker, lust’
Ruc sak ‘dregs’, OM [t]a?dak ‘filth, faeces’, Kharia kɔṭɛ? ‘be dirty’	*(n)zak	AN *[d,D]ak[ih] ‘dirt on skin’
Mon mih ‘body dirt’, Ho homu, Che’ Wong kamah ‘dirty’	*məz	AN *cemeD ‘dirty’
Pacoh kəsaaŋ ‘make a sound’, Sora kədiŋ-ən ‘a drum’, Katu dɨɨŋ ‘play drum’	*(n)zaŋ	AN *genDaŋ ‘drum’
Chrau suk pələy, Khmer (Old) sruk ‘village’, Bonda ḍikh ‘to stay’	*(n)zuk	AN *Duk ‘dwelling place, residence’
Semelai cəruus, Jehai cənrɔs ‘claw (nail)’, PM *krmpuus ‘finger, toe, breadth of finger’	*caḷus, *caḷumpusi	AN *[s]ilu[h] ‘fingernail’
VN cɤp ‘(of heaven) lighten, blink, wink, lightning’, Pacoh picɨɨp ‘go by light of torch’, Thavung ?əcooɔp məlooɔy ‘lightning’	*z[ə]p	AN *han[D]ap ‘flicker’
Bahnar kəsaay? ‘sprinkle, splash water’, Pacoh saay? ‘splash water on self or other’, NK cha? (daak) ‘pond’	*saq(i)	AN *bahaq ‘flood(ed)’
Kharia sɔ?lui, OM sok, Sengoi sok ‘hair’	*suk	AN *buhuk ‘hair’
Bonda ɔŋ, Katu səŋ, Brao ?dɔwŋ ‘hear’	*(n)zəŋiɣ[10]	AN *DeŋeR ‘hear’
Bonda si? ‘fever, pain’, PW *sɨ? ‘pain, disease’, PNB *ji? ‘sick’	*ziq[11], *jiq	AN *pe[d,D]iq ‘hurt, smart, sting’
VN (*zay >) tay ‘at, in’, PM *p()?dəy ‘inside, in, in the middle’, Katu kədaay ‘behind’	*(n)zay	AN *Dayah ‘inland/interior’
Bahnar səlaam ‘in between, in the middle’, Katu (High) jərɨɨh jəruum ‘underneath’, PSB *pəraam ‘inside’	*zaḷəm, *jaḷəm	AN *Dalem ‘inside’
Gutob suku? ‘ladle’, Pacoh sɔɔk ‘dip or scoop food with a spoon’, PW *dok ‘tray’	*(n)zuk	AN *ci(n)Duk ‘ladle, scoop’
Bonda sɔmɔr ‘flame to rise’, PW	*nzaR(i),	AN *damaɣ ‘light, resin, torch’

---

[10] Replaces */zəŋ[əɣ(i)]/ cited in Hayes 1999:24.

[11] Replaces */(n)diq/ cited in Hayes 1997b:24.

*[des] ‘start a fire’, Sengoi dʌr ‘burn, flame up’	*zamaR	
Stieng saw, Kharia yɔ ‘see’, Sengoi tinyaw ‘to watch, look’	*zaw, *ñjaw	AN *tin[D]aw ‘look at closely’
OM kandar ‘wife’, Jeh dri-dri ‘female’, Sedang kodrai ‘female, woman, wife’	*nzar	AN *Darah ‘maiden, virgin’
Katu takɔh ‘grown, old man or woman’, Khmer cas ‘be old, grown up’	*tunqas, *tiqasi	AN *tuqah ‘old’
Mon sapaa ‘flat surface, back (of hand, foot)’, Pearic spal ti: ‘hand’, spal siŋ ‘foot’, Sora sɑ:pɑ:-n ‘ham’	*zapal	AN *DapaN ‘palm, sole’
Katu saal ‘pound rice’, Kharia sɔl ‘mortar for pounding paddy’, Mundari sɛɛl ‘husking hole’	*sa[l̩u]	AN *haluh ‘pestle, pounder’
Rengao həra, PM *(ʔa)da, Kharia gɛṛɛ ‘duck’	*(n)zara	AN *da[r]ah ‘pigeon/dove’
VN ʔdiə, Khmu’ sndɛh ‘dish, plate’, Khmer khdah ‘frying pan, sauce pan’	*nzeh[12]	AN *kenD[ih] ‘pitcher, water jar’
Jehai kəlangis ‘liver’, Semelai gnɔs, PMN *nus ‘heart’	*g[o]si, *ganosi	AN *sa(ŋ)guh ‘pith, sago’
Kharia tɑñ, PW *tañ, Katu taañ ‘weave’	*zali	AN *DanDan ‘plait ropes’
Pearic khsal, OM kyaal, Brou kuyaal ‘wind’	*zal, *ñjal	AN *huDan, *quzan, *quZaN ‘rain’
Souei səŋkaal, Bahnar həkaar ‘skin’, PM *cŋkɔɔr ‘bark of tree’	*sa(n)qaɣ	AN *haka[r] ‘root’
Bahnar sut ‘wipe’, Pearic dus skip, Khmer ṭus ‘rub’, Pacoh chut ‘scrub’	*(n)zoci	AN *kaDus ‘rub, scratch’
Pacoh saat ‘wipe off, rub’, Katu (High) ʔyaat ‘scrub, clean, rub’, MUK cɐt ‘wash clothes’	*zat, *ñjat	AN *DasDas ‘rub off’
Santali bulu, Rengao blu, Nicobar (Central) pulɔ: ‘thigh’	*buḷu	AN *luluD ‘shin’
Sora sərub ‘suck, sip’, Khmer sruup ‘swallow, sip’, Katu kəsruup ‘eat noisily’	*suɣup	AN *hiRup ‘sip’
Juang doko ‘sit’, Khmer duk ‘set (in place), put, place’, Katu (High) dɑɑk ‘put aside for future’	*nzuk	AN *DukDuk ‘sit’
Pacoh cuur ‘sleepy or sad eyes’, Sora kədur ‘snore’, Birhor duṛum ‘sleep’	*(n)zoɣ	AN *tiDuR ‘sleep’

---

[12] Replaces */(kən)[z]eh/ cited in Hayes 1997b:29.

Sora okij ‘a little more’, Katu məcɨy? ‘short time’, Pearic keec ‘small’	*zekiq	AN *Dikiq ‘small, little’
Mundari rowa, Sengoi ruai ‘soul, spirit’, VN (*[h]wa:ś >) vay ‘ancestor’	*r(a,u)wa(s)(i)	AN *ñawah ‘spirit, soul’
Bahnar kəsɔh, OM ksas ‘spit’, Khmer khjaak ‘spit out’	*zaq(i), *jaqi	AN *ludaq ‘spittle’
PM *pruus ‘squirt’, Chrau vruuh ‘squirt, spit’, Khmer bruəs ‘spit onto’	*buɣas(i)	AN *buRah ‘spray, sprinkle’
Chrau sɤɤr ‘take out, up’, Sora ub-yeer ‘rise up’, Pacoh yoor ‘get up, arise’	*zeɣ[13], *ñjeɣ	AN *DiRih ‘stand (up)’
PM *?buh ‘boil’, Khmer buh ‘boil, seethe, bubble’, Pacoh boh ‘put into fire to roast’	*(m)bus	AN *sebuh ‘develop steam’
Khmer phsaar ‘join (two members) with gum, glue or the like’, Bahnar jar ‘pitch of tree’, Khmer jaar ‘sap, resin’	*zər, *(pi)jər	AN *pizer ‘to stick’
Pearic sro(:)t ‘undress oneself’, VN thɔt ‘pull in (one's stomach)’, Sora ruj ‘pluck, pull out’	*sarut(i)	AN *hurut ‘stroke’
Kharia daw ‘big flat sickle’, Pearic (kɔ)ndiəw ‘sickle’, Khmer ṭaaw ‘sword’	*(n)zaw	AN *maŋ[d,D,j]aw ‘war sword’
Juang goneh ‘tooth’, PM *gnis ‘canine tooth’	*g[a]nis	AN *gigih ‘tooth’
PW *rɨs ‘turtle’, PM *kn?duh ‘tortoise’	*(kən)[ḷ,r]us	AN *peñuh ‘turtle’
Pacoh ŋeaih ‘count’, OM ŋus, Lamet ŋɔɔs ‘price’	*(i,u)ŋkasi	AN *zaŋkah ‘unit of measure’
Sora s?i:ŋ-ən ‘a house, an abode, a room’, Chrau ?ding, PW *[ntiŋ] ‘wall’	*(n)zeŋ	AN *[d,D]iŋ[d,D]iŋ ‘wall’
Khmu’ ?om ‘water’, Khasi sum ‘bathe’, Gutob dum ‘drown’	*[?]om, *(n)zom	AN *[d,D]anum ‘water/fresh water’
PW *pes, Katu (*piś > *piyh >) piih ‘sweep’	*(tam)pis	AN *ta(m)pih ‘winnow’
Brou sɔɔy ‘tail’, Pacoh ndɔɔy ‘little finger/toe’, VN (*jɔɔy >) zɔy ‘worm, larva’	*u(n)zay	AN *huDay ‘worm’

---

[13] Replaces */(z)yeɣ(i)/ cited in Hayes 1999:26.

Comments on Hayes „The Austric Denti-alveolar Sibilants“

By Václav Blažek
Ryneček 148
Příbram III
Czech Republic
blazek@phil.muni.cz

Reading this stimulating article, in which the author tries to demonstrate regular correspondences among some Austronesian (AN) and Austroasiatic (AA) consonants, I find the weakest point to be in the Austroasiatic [pre]reconstructions. That is why I check the accessible partial reconstructions of the lower units of the Austroasiatic family, to judge their compatibility with the impressionistic AA reconstructions of the author.

pMK _*ris_ "root" (Peiros) > pBahnar _*riəh_ (Sidwell); Khmer _rɨh_; Mon _ruih_; pKatuic _*rɨejh_ (Peiros); pWaic _*rɛs_ (Diffloth); pPearic _*re:s_; pVM _*re:lh_ (Sokolovskaja) - see Sidwell 1998: # 332. Cf. also Munda: Mundari _redʔ_ "root"; Central Nicobar _yiah_ id., and Asli: Semang of North Perak _yaes_ "root". I prefer the comparison with AN _*ʔuɣat_ (Dempwolff) = _*uR_$_2$_at_ (Biggs) "sinew, vein" and further pAinu _*rit_ "root" (Vovin), cf. Kuril (Voznesenskij) _ryt_ "tendon" (see Bengtson & Blažek 2000: # 66).

Katu _saak_ "corpse" and _ʔəcak_ "body" (pKatuic _*ʔə|caʔ_ after Peiros) represent different etyma, cf. the other AA cognates: pBahnar _*kə|ja:k_ "ghost of dead", Vn _xác_ "corpse"; Khmer _biʔsac_ "devil" vs. pBahnar _*cak_ "body, name" (Sidwell 1998: ## 47, 171 respectively).

pWaic _*som_ "night" is comparable with Mon _samò_ "evening" and further probably with pSBahnar _*mhʌ:_ id. and Palaung _hmȭ_ "night, evening" (Efimov 1990: 113).

pMK _*suək_ ~ _*sək_ "hair" continues in pBahnar _*šok_ ~ _*šɔk_; Khmer _suk_; pVM _*sok_; pPearic _*suk_; pKatuic _*sɔʔ_; pWaic _*hɨk_; Mon _sok_ (Sidwell 1998: # 460). Following Diffloth, Peiros (1998: 163) also compares the AA data with AN _*buSek_ ~ _*buSuk_. Blust's reconstruction better agrees with the AA counterparts.

pKatuic _*saŋ_ "to hear" corresponds to pPearic _*saŋ_; Car Nicobar _haŋ_ id. and pBahnar _*kə|təŋ_ ~ _kə|taŋ_ "to hear, listen" (Sidwell 1998: #764).

pBahnar _*ɟi:ʔ_ "sick" pWaic _*sɨʔ_ and Munda: Bonda _siʔ_ "fever, pain" correspond to Khmer _chɨ:_ "sick" and Car Nicobar _cək_ "pain" (Sidwell 1998: # 267).

Bahnar _səlaam_ "in between" must be derived from the primary root _lam_, _ləm_ "in, inside" < pBahnar _*lam_ ~ _gə|lam_ (Sidwell 1998: # 185 compares it with pChamic _*dalăm_ "inside, deep," which continues AN **Daləm* "inside, depth" in the reconstruction by Dempwolff). The comparison of the root elements, i.e. **lam // *ləm,* is quite acceptable.

Sedang *kədraj* and Jeh *dri:dri:* "female" reflects pBahnar **kənri:* id.; this reconstruction is compatible with pWaic **krih* "unmarried woman" (Sidwell 1998: # 263) and OMon *kandar*, *kindar* "wife" (Diffloth 1984: 116 reconstructs pMon **kmdar*).

pMon **(ʔa)daa* "duck" (Diffloth 1984: 68) corresponds to pBahnar **tə|da:* id.; Khmer *tiə* id.; pKatuic **ʔədia* ~ **ɣədia* id. (Sidwell 1998: # 4).

pMnong **nus* "heart" reflects pBahnar **(ʔ)nu:š*. It is compatible with Asli: Semelai *gnɔs* id., while the relationship of Jehai *kəlangis* "liver" remains [for me] uncertain. The comparison with AN **sa(ŋ)guh* "pith, sago" (Lopez) remains problematic, especially for semantic reasons.

We should also consider pMK **kja:l* "wind" (Peiros 1998: 159, who adds pMunda **kojo* and pMY **C uL* "wind") > pBahnar **kər|ja:l* ; Khmer *khjal*; pVM **kjɔʔ*; pKatuic [*k/g*]əʔ*ja:l*; pWaic **kɨr* id.; Car Nicobar *kiso:l cin* "I blow"; Sidwell 1998: # 56 who also mentions ST **qh*^{w}*iał* "wind"). Concerning AN, the reconstruction **quZ*$_{1}$*ále* "rain" proposed by Dyen & McFarland is most compatible with the AA data.

pBahnar **ʔə|ka:r* "skin" and pKatuic **həŋkar* id. plus pMon **cŋkɔɔr* "bark of tree" reflect pMK **(CəN)kVr* (Peiros apud Sidwell 1998: # 95). The comparison with AN **haka*[*r*] "root" (Lopez) = **w*$_{4}$*aka*[*rR*] (Dyen & McFarland) is perhaps possible although the semantic difference is not trivial.

Khmu *ʔom* "water, river" is more compatible with Khasi *u:m* id. and pBahnar **ʔum* ~ **hum* "to bathe", pWaic **rʔom* "water" & **hɨm* "to bathe"; pPearic **ʔu:m* "to bathe" (Sidwell 1998: ## 600, 604) than with Khasi *sum* "to bathe" (prefix ?). Peiros 1998: 143 also adds pMY **ʔvom* "water". On the other hand, AN ***[*dD*]*anum* "(fresh) water" is probably more comparable with pBahnar **ʔno:m* "to urinate"; Khmer *naom*; pWaic **( )nɨm* id. (Sidwell 1998: # 488); Asli: Sakai *kenám* id. and further pYao **ʔnam* "cold of water" and Kadai **R-nam* "water" (Peiros 1998: 143) and even Ainu *nam* "fresh or cool (as fresh water), cold as water or one's feet hands", *nam wakka* "fresh or cool water" (Bengtson & Blažek 2000: # 78).

Bru *sɔɔy* "tail" is perhaps better derivable from pMK **Cəduj* "tail"; cf. also pMY **toi* id.; AN **hudi* "buttocks" (Peiros 1998: 157, 165).

Summing up, I find the arguments for the postulated AA **z* and its correspondence to AN **D* unconvincing. From Hayes' comparanda studied here, I am able to accept the entries "hair", "wind / rain", "bark of tree / root", plus the comparisons proposed or supplemented by other authors, e.g. "root", "water", "tail". Without careful AA reconstructions based on the partial reconstructions of the daughter's protolanguages, the AA-AN comparisons remain only speculative.

Abbreviations

MK Mon-Khmer, MY Miao-Yao, O Old, p proto-, VM Viet-Muong, Vn Vietnamese.

References

Bengtson, John D., & Václav Blažek. 2000: Lexical Parallels Between Ainu and Austric and Their Implications. *Archív orientálni* 68/2.

Diffloth, Gérard, 1984: *Dvaravati Old Mon Language and Nyah Kur.* Chulalongkorn: University Printing House.

Efimov, Aleksandr Jurjevič, 1990: *Istoričeskaja fonologija južnobahnaričeskix jazykov.* Moskva: Nauka.

Peiros, Ilia, 1998: *Comparative Linguistics in Southeast Asia.* Canberra: The Australian National University.

Sidwell, Paul, 1998: *Proto-Bahnar comparative-historical phonology and comparative etymological lexicon.* Melbourne: Diss.

Wurm, S.A. & Wilson, B. 1975: *English Finderlist of Reconstructions in Austronesian languages (post-Brandstetter).* Canberra: The National University.

Comments on Hayes,
"The Austric Denti-alveolar Sibilants"

By Robert Blust
University of Hawaii

Anyone familiar with the workings of science in any of its varied manifestations will sense that something is fundamentally wrong with a empirical proposal when its author speaks with unshakeable confidence of 'proof,' much less 'irrefutable proof.' Putting aside questionable cases such as string theory in cosmology, scientific hypotheses are not, and never have been, products of the deductive logic which governs mathemathical proofs. Rather, they are products of an inductive logic which is forever open to falsification, but never to confirmation. There is an important trade-off here: while mathematics deals with certainty, it does not deal with the real world, nor can new information enter the chain of inference once it is started. In short, mathematics is concerned with the internal logical consistency of symbolic systems. Consequently it offers few surprises. Science is messier but, like real life, is full of surprises. In place of certainty, we must settle for probability as measured by plausibility, economy, independence of evidentiary support and the like. But, to partially compensate for this perennial incompleteness, science can make testable statements about the real world. The inductive logic of science generalizes based on limited samples, makes predictions which extend beyond the sample, and provides a clear basis for falsifying the predictions.

The present work is not a piece of science. It makes no use of the Comparative Method of historical linguistics, but in classic amateur fashion simply searches for general phonetic and semantic similarities which it then pairs in proposed 'cognate sets.' In scanning the data given here, one searches in vain for recurrent sound correspondences in plausible comparisons. Genuine evidence of genetic relationship can often (although not always) be shown with only limited sets of data. Consider the following comparisons between Malay and Hawaiian, two Austronesian languages which have been separated for about 5,000 years. To simplify the argument, we will assume that correspondences of identical phonemes require no further comment:

NO.	MALAY		HAWAIIAN	ENGLISH
1)	**mata**	:	**maka**	'eye'
2)	**kutu**	:	**'uku**	'louse'
3)	**ikan**	:	**i'a**	'fish'
4)	**langit**	:	**lani**	'sky'
5)	**tangis**	:	**kani**	'cry; wail'

In 1) the comparison can be considered valid if we are able to establish that Malay /t/ : Hawaiian /k/ is a recurrent sound correspondence. Support for this hypothesis comes from 2), provided that we also assume that Malay /k/ : Hawaiian /'/ is a recurrent sound correspondence. Support for this second hypothesis comes from 3), provided that we also assume that the correspondence of Malay final consonants to final zero in Hawaiian is recurrent. Support for this this third hypothesis comes from 4) and 5), and from the general observation that Hawaiian has no final

consonants, whereas Malay permits both final consonants and final vowels.Finally, the correspondence of Malay /ng/ (velar nasal) to Hawaiian /n/ is seen in each of the last two examples. With just five comparisons, then, we are able to demonstrate that the following correspondences are recurrent: 1. /t/ : /k/ (examples 1, 2 and 5), 2. /k/ : /'/ (examples 2 and 3), 3. final consonant to zero (examples 3-5), 4. /ng/ to /n/ (examples 4 and 5). Chance could produce any one of the above comparisons, but it certainly could not produce the tight interconnection of sound correspondences in words of essentially identical meaning in all five (and, of course, many more).

It is known that I have come down on the side of an Austric hypothesis that links the Austroasiatic languages and Austronesian in a distant genetic relationship, probably on the order of 9,000 years separation time. (See Blust 1996 for the reasoning which leads to this estimate). But the evidence for Austric shows a striking imbalance: key elements of the morphological core seem to have been preserved long after lexical replacement has virtually eliminated evidence of recurrent sound changes in cognate bases. It would be a welcome addition to the argument if Hayes could demonstrate, through application of the Comparative Method, that Austroasiatic and Austronesian do, indeed, share lexical similarities which could not plausibly be explained as products of chance or borrowing. But he has not done that in this paper. What we see is a grasping at straws, not evidence of recurrent sound correspondences.

To compound the problems of taking these comparisons seriously, many of the Austronesian reconstructions are erroneous. Hayes evidently has relied primarily on Dempwolff (1938), with a few concessions to typographical changes made during the 1940's by Isidore Dyen. But Dyen's (1953) fundamental revision of Dempwolff's treatment of the 'laryngeals' - though published nearly half a century ago - is totally ignored. To take just the first three Austronesian reconstructions, the first should be *biRaq; the second is not regarded as an early Austronesian form, but rather a product of borrowing from Malay within western Indonesia; and the third is *Sawak. In other words, Hayes is using *h for at least two distinct Proto-Austronesian phonemes which have quite different reflexes in a number of languages (*q is reflected as a pharyngeal stop in several Formosan languages, and as glottal stop, /k/, /h/ or zero outside Taiwan, whereas *S is reflected as a sibilant in most Formosan languages, but as /h/ or zero outside Taiwan). From 1990-1995 the writer was funded by the National Science Foundation to produce a new comparative dictionary of the Austronesian languages, the *Austronesian Comparative Dictionary* (ACD). Although it presently is only about 25% complete, the ACD contains over 2,000 printed pages of reconstructions on various explicitly marked levels (Proto-Austronesian, Proto-Malayo-Polynesian, Proto-Oceanic, etc.), together with supporting evidence from around 150 languages. Some of this material was previously published in a series of articles totalling over 500 pages in the journal *Oceanic Linguistics*. Access to the ACD has periodically been given over the internet, and, in the modern 'information age', it is surprising to see an author ignore virtually all relevant source material less than 50 years old.

Many of the comparisons given by Hayes arbitrarily choose one or the other syllable for the proposed match, while ignoring the other, as with Khmer /ṛas/ 'root' (presumably compared with the last syllable of PAN *biRaq), but Ruc /sak/ 'dregs' (presumably compared with the first syllable of PAN *daki 'dirt on the skin'). Some comparisons, such as Proto-Waic *rɨs 'turtle' with PAN *peñu 'green turtle', or Vietnamese /vay/ 'ancestor' with PAN *ñawa 'breath', are strange, to say the least - the product of erroneous etymologies combined with a determination to 'prove' a relationship. One looks in vain for the kind of accountability that one finds in responsible applications of the Comparative Method. In short, this is not good science, and it does not advance our knowledge of the linguistic history of Southeast Asia and the Pacific.

REFERENCES

Blust, Robert. 1990, 1991, 1992, 1993, 1994, 1995. *Austronesian Comparative Dictionary*. [Work in progress].

Blust, Robert. 1996. Beyond the Austronesian homeland: the Austric hypothesis and its implications for archaeology. In Ward H. Goodenough, ed., Prehistoric settlement of the Pacific. *Transactions of the American Philosophical Society* 86.5:117-140. Philadelphia: American Philosophical Society.

Dempwolff, Otto. 1938. *Vergleichende Lautlehre des austronesischenWortschatzes*. Vol. 3: Austronesisches Wörterverzeichnis. Berlin:Reimer.

Dyen, Isidore. 1953. The Proto-Malayo-Polynesian laryngeals. Baltimore: Linguistic Society of America.

Four Austric Theories

By George van Driem

Leiden University

Wilhelm Schmidt is the father of Austric, a theory which posits a language family consisting of Austroasiatic and Austronesian. Frans Kuiper was perhaps the first scholar after Schmidt to adduce a good number of possible Austric correspondences, some of which show stunning semantic and formal correspondence (1948). In the *Mother Tongue* of October 1999, I wrote that Frans Kuiper's comparison of Malay prefixes with a hypothetical Austric source language for the early loans in the R̥gveda would be a fanciful exercise unless the correctness of the Austric theory was presumed. I should, of course, have written that Kuiper's comparison was meaningful because he presumed the correctness of the Austric theory. Moreover, Kuiper was the first to address the question of identifying possible Austroasiatic prefixes in the unindentified loan layer in the R̥gveda. The hypothesis formulated by Kuiper was that the early loan layer exhibited elements which could be identified as ancient Austroasiatic prefixes, only relicts of which could be found in Munda but many of which were still found intact in Malay.

Wilhelm Schmidt's Austric was a macrofamily, which was later even to include Japanese as a predominantly Austric 'Mischsprache' consisting of an 'austroasiatische' and an 'ural-altaische' layer (1906, 1930). There are also other versions of the Austric theory. The oldest alternative version is nearly as old as Schmidt's Austric. What were seen as correspondences between Siamese, Chinese, Burmese, Tibetan and Malay led both August Conrady (1916, 1922) and Karl Wulff (1934) to espouse the hypothesis of a genetic relationship between Indo-Chinese and Wilhelm Schmidt's Austric, whereby Indo-Chinese consisted of Daic plus Tibeto-Burman. In those days, Chinese was seen as closely related to Siamese, rather than as a constituent branch of Tibeto-Burman (cf. van Driem 1997, 1999a). I call this theory mega-Austric. It proposes a superfamily consisting of Austroasiatic, Austronesian, Daic and Tibeto-Burman. There exists yet another version of Austric to which I gave the name 'Greater Austric' several years ago (1998). This hypothetical construct comprises Austroasiatic, Austronesian, Daic and Hmong-Mien and is therefore more comprehensive than Schmidt's Austric, yet not quite as inclusive as Conrady's mega-Austric. Greater Austric unites the Austric and Austro-Tai theories. Robert Blust entertained a version of the Greater Austric hypothesis when he proposed that Austroasiatic might represent one trunk of 'the Austric superfamily' with Austro-Tai (i.e. Daic plus Austronesian) making up the other trunk. (1996).

La Vaughn Hayes is a proponent of Schmidt's Austric, but does not reject the possibility of a genetic relationship of Austric with Daic and Hmong-Mien. Though not much evidence has accrued for either Greater Austric or mega-Austric since August Conrady and Karl Wulff had a look at it, Hayes claims that for Austric proper there is 'massive evidence of a shared core vocabulary, only a small part of which' he adduced in his first article (1992: 174). In his and in his most recent article he even sets up a sound law involving regular correspondences between Austroasiatic and Austronesian for reflexes of the reconstructed Austric sibilants *s and *z, which he argues constitutes 'irrefutable proof' for the reality of Austric. This is evidence which deserves to be taken seriously, especially in view of the great time depth which is usually assumed for Austric. It is to be hoped for that young scholars will pick up the gauntlet which Hayes has cast down, and that new initiatives will be undertaken to describe the many hitherto undescribed Austroasiatic languages in rigorous detail.

Kuiper felt that 'the relatively small number of words which Austronesian has in common with Austroasiatic is not, accordingly, sufficient proof in itself to assume that both branches have sprung from one parent language' (1948: 380), and Kuiper therefore looked for additional evidence in the form of morphological correspondences between Malay and Munda. Wilhelm Schmidt first presented morphological evidence for his Austric hypothesis in 1906, much of which was drawn from Nicobarese, and Lawry Reid has continued this tradition (1994). I discussed several problems with the four proposed morphological parallels which constitute all the evidence in my essay on the Austroasiatic Indus Theory in the last *Mother Tongue*. The purported resemblances lose much of their force when the facts are viewed in fuller comparative context. It is sobering to recall that the existence of purported reflexes in Old Japanese of the Malayo-Polynesian infix *<-um-> had already been adduced early in the 20th century as evidence for what is now called the Austro-Japanese theory, or the 'Nippon-Malay-Polynesian' language family as it was first called by Dirk van Hinloopen Labberton, the professor of Dutch in Tokyo who originally proposed the theory in 1924.

Before most of Hayes' work had been published, Gérard Diffloth had examined the scanty lexical evidence for Austric thus far that can bear up to scrutiny, and established that the lexical evidence was on the whole negative. Yet there seems to be sufficient evidence, both material and circumstantial, to take the theory seriously, whether Schmidt's Austric (i.e. Austroasiatic and Austronesian), Blust's Greater Austric (Austroasiatic, Austronesian, Daic and possibly Hmong-Mien) and Conrady's mega-Austric (i.e. Austroasiatic, Austronesian, Tibeto-Burman, Daic and presumably also Hmong-Mien). The lexical and morphological evidence for Austric may also make sense, and perhaps even more so, if a fourth and totally novel variation on the Austric theme is assumed. This new, fourth version of the Austric theory was proposed by Frederik Kortlandt and is a radical departure from the conventional view that Austroasiatic shows too much internal diversity to have formed at a time depth more shallow than three millennia. This version of the Austric theory entails that Austroasiatic is a branch of Malayo-Polynesian, which makes the time depth of Austroasiatic shallower than that of Malayo-Polynesian and much shallower therefore than that of Austronesian as a whole. This fourth version of Austric makes sense if it is presumed that the intrusive proto-Austroasiatic branch of Austronesian was subjected to rapid interference through shift and bilingualism after migrating to the Southeast Asian mainland in much the same way as ancient Austronesian languages were when they were transplanted to New Guinea. The interference and intense contact situations presumed for Austroasiatic are precisely what would be expected if an intrusive seafaring populace were to settle the already populated Southeast Asian mainland. As a result, Austroasiatic would show great internal diversity and appear an order of magnitude older than, say, the Oceanic languages of the Pacific, the spread of which represents the opposite situation of pioneers colonizing virgin, previously uninhabited islands rather than intrusive groups having to assimilate a heterogeneous collection of resident mainland populations.

Stanley Starosta showed that prefixing as a morphological process began in western Formosa, whence it spread via the northeast to the southwest of the island, and thence via the Philippines to become a full-fledged system in Malayo-Polynesian. Kortlandt's hypothesis that Austroasiatic is an offshoot of ancient Malayo-Polynesian therefore makes sense of the morphological correspondences adduced by Lawry Reid (1994). This morphological evidence falls into place when viewed in light of the relative chronology, elucidated in Starosta's work, of the development and grammaticalization of affixal processes in Austronesian. Kortlandt's Austric also makes sense of the lexical correspondences between Munda and Malay adduced by Frans Kuiper, which involve many items with strikingly specific formal and semantic correspondence, e.g. Santal *gạvić,* Mundari *gāui,* Kharia *gouj* 'beckon with the hand' vs. Malay *gamit* 'touch slightly with the finger in order to give a hint or draw attention' (1948: 377). Kortlandt's version of Austric would explain why Kuiper was able to adduce correspondences at this level, whereas the result of Diffloth's lexical comparison was largely negative, as it was based on the assumption that Austroasiatic and Austronesian were coordinate nodes. In other words, the hypothesis that Austroasiatic is an ancient offshoot of Malayo-Polynesian and not a coordinate node with Austronesian makes sense of the lexical correspondences which are in evidence between Malayo-Polynesian and Austroasiatic in contrast to the relative paucity of lexical correspondences between Austroasiatic and Austronesian.

Kortlandt's Austric also presumes a far more likely ancient prehistoric migration than Schmidt's Austric because it involves a maritime migration to the Southeast Asian mainland from insular Southeast Asia, whereas Schmidt's Austric theory necessitates a migration overland from a putative *Urheimat* to both Formosa, the recognized Austronesian homeland, and to the area surrounding the Bay of Bengal, the Austroasiatic centre of gravity. On the other hand, the Southeast Asian mainland is an obvious and probable destination for a seafaring race such as the Austronesians, who managed to colonize such out-of-the-way and improbable destinations as Madagascar, Easter Island and Hawai'i. Roger Blench has proposed that archaeological evidence which could be interpreted to support the hypothesis that Austroasiatic is a branch of Austronesian would be the connexion between the Lapita ware of Formosa, the Philippines and eastern Indonesia and the red slipped ware tradition of mainland Southeast Asia. This interpretation fits chronologically because the connexion is quite late, dating from between the middle and the end of the second millennium BC. The Malayo-Polynesians had fanned out from their Formosan homeland long before then, and mainland Southeast Asia was most certainly already inhabited by other, non-Austronesian peoples, for the neolithic assemblages in Thailand and Vietnam are about a millennium older than those of Indonesia. The appearance of red slipping on the Southeast Asian mainland at this time is compatible with the hypothesis of an intrusive Austronesian population ancestral to Austroasiatic because the red slipped ware found in Thailand, Vietnam and Malaya is associated with cord marked pottery styles which do not occur in insular Southeast Asia, and there is little resemblance between the earliest pottery of Thailand and Vietnam and that of the Philippines and Indonesia.

Another tell-tale sign of a linguistic intrusion is the manifest racial difference between Munda speakers on one hand and the Nicobarese, the Khasi and the speakers of Mon-Khmer languages on the other hand. This physical difference could be accounted for by assuming that the Munda are the descendants of a pre-Austroasiatic group which learnt Austroasiatic. Kortlandt's version of Austroasiatic as an offshoot of Austronesian is compatible with Robert von Heine-Geldern's theory of an Austric homeland in mainland Southeast Asia, though there is a difference of time depth. Kortlandt's theory supposes a time depth of just over three millennia and can be tentatively identified with the intrusive appearance of red slipped ware in Thailand and Vietnam, whereas Wilhelm Schmidt's Austric consisting of Austroasiatic and Austronesian as two coordinate nodes would have to be of far greater antiquity and has therefore been identified with the mesolithic Hòabìnhian technocomplex. If we assume Kortlandt's scenario, the arrival of various Tibeto-Burman groups in northwestern India from Sìchuān may have been the disruptive force which drove the linguistic ancestors of the Munda further west deeper into India, whereas the arrival of the Pyu in the Irrawaddy basin is what split up Mon-Khmer. These competing theories and scenarios are discussed in greater detail in Chapter Two of my forthcoming handbook on the greater Himalayan region, entitled *Languages of the Himalayas*. The linguistic evidence holds primacy above the archaeological evidence because obviously only language can provide incontrovertible evidence of a linguistic intrusion and the spread of a language family. The resolution of the Austric problem is pivotal to our understanding of Asian prehistory. Yet at this point the epistemological basis for Austric is still meagre as far as language families go, let alone for Greater Austric or mega-Austric. For this reason, the comparative investigations conducted by La Vaughn Hayes are of great value in this much neglected field. The most urgent task before us, however, is the detailed documentation of the many hitherto undescribed or only partially described Austroasiatic languages, most of which are currently endangered with extinction.

References

Blust, Robert. 1996. 'Beyond the Austronesian homeland: The Austric hypothesis and its implications for archaeology', pp. 117-160 in Ward H. Goodenough, ed., *Prehistoric Settlement of the Pacific* (Transactions of the American Philosophical Society, Vol. 86, Pt. 5). Philadelphia: American Philosophical Society.

Conrady, August. 1916. 'Eine merkwürdige Beziehung zwischen den austrischen und den indochinesischen Sprachen', pp. 475-504 in *Aufsätze zur Kultur- und Sprachgeschichte vornehmlich des Orients: Ernst Kuhn zum 70. Geburtstage am 7. Februar 1916 gewidmet von Freunden und Schülern.* München: Verlag von M. & H. Marcus.

Conrady, August. 1922. 'Neue austrisch-indochinesische Parallelen', pp. 23-66 in *Asia Major: Hirth Anniversary Volume*. London: Robsthan and Company.

van Driem, George. 1997. 'Sino-Bodic', *Bulletin of the School of Oriental and African Studies,* 60 (3): 455-488.

van Driem, George. 1998. 'Neolithic correlates of ancient Tibeto-Burman migrations', pp. 67-102 in Roger Blench and Matthew Spriggs, eds., *Archaeology and Language II.* London: Routledge.

van Driem, George. 1999a. 'A new theory on the origin of Chinese', pp. 43-58 in Peter Bellwood and Ian Lilley, eds., *Bulletin of the Indo-Pacific Prehistory Association Bulletin, 18 (Indo-Pacific Prehistory: The Melaka Papers, Vol. 2).* Canberra: Australian National University.

van Driem, George. 1999b. 'On the Austroasiatic Indus Theory', *Mother Tongue*, Special Issue (October 1999): 75-83.

van Driem, George. [in press]. *Languages of the Himalayas.* Leiden: Koninklijke Brill.

Hayes, La Vaughn H. 1992. 'On the track of Austric: Part I', *Mon-Khmer Studies,* XXI: 143-177.

Hayes, La Vaughn. 1997a. 'Another look at final spirants in Mon-Khmer', *Mon-Khmer Studies,* XXVI: 51-64.

Hayes, La Vaughn. 1997b. 'On the track of Austric, Part II: Consonant mutation in early Austroasiatic', *Mon-Khmer Studies*, XXVII: 13-41.

Hayes, La Vaughn. 1999. 'On the track of Austric, Part III: Basic vocabulary correspondence', *Mon-Khmer Studies,* XXIX: 1-34.

Hayes, La Vaughn. 2000. 'The Austric denti-alveolar sibilants', *Mother Tongue.*

Kuiper, Franciscus Bernardus Jacobus. 1948b. 'Munda and Indonesian', pp. 372-401 in *Orientalia Neerlandica: A Volume of Oriental Studies, published under Auspices of the Netherlands' Oriental Society (Oostersch Genootschap in Nederland) on the Occasion of the Twenty-Fifth Anniversary of Its Foundation, May 8th, 1945.* Leiden: A.W. Sijthoffs Uitgeversmaatschappij.

Reid, Lawrence A. 1994. 'Morphological evidence for Austric', *Oceanic Linguistics,* 33 (2): 323-344.

Schmidt, Wilhelm. 1906. 'Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens', *Archiv für Anthropologie,* Neue Folge, V: 59-109.

Schmidt, Wilhelm. 1930. 'Die Beziehungen der austrischen Sprachen zum Japanischen', *Wiener Beiträge zur Kulturgeschichte und Linguistik, Veröffentlichungen des Institutes für Völkerkunde an der Universität Wien,* I: 239-251.

Wulff, Karl. 1934. *Chinesisch und Tai: Sprachvergleichende Untersuchungen* (Det Kongelige Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser, XX, 3). Copenhagen: Levin & Munksgaard.

LaVaughn Hayes and Robert Blust Discuss Austric

Comments by Hal Fleming

Gloucester, Mass.; Past ASLIP President

My most noteworthy thing to say is that I am woefully unqualified to comment on the details of a discussion about Austric. Both LV Hayes and Robert Blust know far more about the kinds of etymologies which have been presented and about the history of the whole discussion. I may contribute only two things: (a) as a kind of globalist I am experienced at looking at the 'proofs' (or evidence), usually lists of etymologies which are presented to support hypotheses of genetic relationship; and (b), in the same vein, I am sensitive to the kinds of arguments made by protagonists and antagonists.

It is important to mention that the question of authorship is not involved here, nor is the argument about the existence of a new phylum. They both agree that Austric existed, that Pater [Wilhelm] Schmidt was the author of it, and that our late revered colleague Paul Benedict had once agreed to Austric and then later argued against it on the grounds that the evidence was solely morphological, the lexical evidence being rendered irrelevant or inconclusive because of massive borrowings. I would add that Joseph Greenberg, in his global taxonomy survey in 1954, had agreed with Schmidt's Austric, which position he re-asserted in 1987 for Ruhlen's *Guide*. Gérard Diffloth, expert on Austro-Asiatic, also agreed on Austric, although he was stymied by Benedict's opposition.

For other non-Southeast Asianists, especially Africanists, Nostraticists, and Americanists, the scope of the discussion will be enlightening. The Austric hypothesis joins two very different entities together, namely, (a) Austro-Asiatic, a set of quite dissimilar languages on the continent from central India to southeast China, and down into the Nicobar islands, and (b) Austronesian, a huge set of quite similar languages - roughly 1000 of them*(1)* - spread from Taiwan throughout the Pacific area, except for Australia, highland New Guinea, and a few islands in Melanesia. Not only had a reflux of Austronesian brought a few of its languages back to the mainland (e.g., Cham in Viet Nam), but also Austronesian had reached across the Indian Ocean to the East African coast, where cultural traits can still be found, to the Comoros where more of that is found, to Madagascar whose native languages are Austronesian. Despite its great size, Austronesian gets compared with Indo-European and Bantu in apparent age and relatively slight diversity. Like the other two, it is a fairly obvious phylum and one that was discovered even earlier than Indo-European, at least in part. That was first in 1603 and a much fuller version by 1706 (Ruhlen, p.161). Bantu had been known and was being related by some workers to various members of what became the Niger-Congo phylum by the mid 19th century.

1. Ruhlen (p.338) lists 959 Austronesian languages which he too accepts as conjoined with 155 Austroasiatic languages, which would total 1114, but his Austric also includes 57 Daic [Kadai] and 4 Miao-Yao [Hmong-Mien] languages for a grand total of 1175 Austric languages. Neither Blust nor Hayes enrolls Daic or Miao-Yao in Austric. Diamond (2000:709) lists about 1200 Austronesian languages, this presumably on the advice of Robert Blust. He also believes that there are about 6000 languages in the whole world, versus Ruhlen's 5000.

Austro-Asiatic, with fewer languages but deeper diversity, was more difficult. Still it has been accepted generally for most of the 20th century. Its ties to Austronesian are deeper still, and its retrievable common lexicon with Austronesian is more difficult to find. Given Paul Benedict's statement that there wasn't any common lexicon, one can assume that the proposed Austric must have a time depth rather like Afrasian or possibly Nostratic or Amerind. That suggests 10kya or more, by a relatively cautious estimate.

We really have to discuss only three things, namely, the quality of Hayesian etymologies, the quality of his argument of "Irrefutable Proof," and the quality or aptness of Blust's criticism.

To begin with, Hayes proposes a number of "similarities" which he believes are cognates which show regular sound correspondences. Some of the proposed etymologies are quite good and convincing. Some are a bit harder to take. And some are somewhat tortured, requiring one to believe that a number of optional shapes (indicated by those irritating parentheses which Nostraticists use so much) add up to cognation. Most of the exercise is conducted with starred forms, as Guthrie would call them, which are always problematic and especially so when the parentheses say in effect that the form is either this or that or maybe a third. One might call this "opportunistic reconstruction," in that one of the options in the parentheses might on some occasion correspond to a form in another language. Hayes also indulges systematically in "implicit segmentation," i.e., we have to believe that part of one word is cognate with all of another word. Thus, if we compared *turn* with *return* we would treat the *re-* of *return* as a prefix of some sort and relate the two *turns* as cognate. Well and good if we know that *re-* is a prefix that can go with a verb like *turn.(2)* Sometimes it is not, as in *rest* or *resting* which are not cognate with 'sit' or 'sting', i.e., they do not segment to 're-sit' or 're-sting'. Nor can we get verbs *–bel* or *–pel* from *rebel* or *repel* - in English - even though their Latin originals did use *re-*. Ditto *regular, region, reign, religion,* etc.

Yes, but we have to ask the 'so what?' question. This is precisely what Trask failed to do with Bengtson [MT I]. Indeed, it is a question which he ignored like a true cavalier. Hayes has a number of etymologies, pairs and/or triplets, which are convincing, are not tortured, and are probably true. We ought not throw out the baby, yet keep the bath water so that we may scorn it. I do not know if the number of 'good etymologies' reaches 35 or not, but certainly Hayes gets close to that by my estimate of baby parts, not bath water.*(3)*

On the second question, the quality of Hayes' argument of 'Irrefutable Proof', our reaction has to be on two levels. On the so-called higher level, we must agree with Blust

---

2. This is the sin that our old friend [R.L.] Trask accused John Bengtson of so frequently in their exchanges on Basque and its Caucasic relatives. And Trask was often correct, and John did render his own argument more difficult to accept because he did not clarify for the reader what assumptions he was making. Trask, of course, turned out to be wrong very often because he was so rigid.

3. Ives Goddard, a well-known critic of Greenberg's Amerind book, said at the famous Stanford conference in 1987 that he would believe any two or more languages were related if he or someone else could find thirty-five good cognate sets. More recently his colleagues have partly gotten around that proposed requirement for genetic relation by eliminating most etymologies, or as many as possible, by claiming that they were borrowings. Indeed what they are getting away with, in so much of what they call borrowing, is probably not justified. The most extraordinary effort in this tradition, however, was the one Wm. Jacobsen made in MT-I to discredit all of Bengtson's etymologies (that remained after Trask's critique) by calling them borrowings from Latin. Most were in basic vocabulary. The reader is invited to review MT-I to see a truly noble effort to spare the world the truth about Basque.

that Hayes seems not to understand what science really does. On the lower or more detailed level, we have to agree with Hayes that the kinds of correspondences he shows - and their historical distributions - argue quite powerfully for genetic connection. We must suggest to Hayes that he read *Mother Tongue* once in a while, especially Greenberg's article on 'Proof' in an earlier issue of this Journal [MT-I, 1995]. Hayes seems to have a typical linguist's notion that mathematical concepts are overwhelmingly important and crucial in linguistics. But we must suggest to Blust that he cease berating Hayes for his lack of sophistication in science, and start paying attention to the serious evidence that Hayes has produced, even while tripping over his scientific misconceptions. One cannot escape the suspicion that Blust has found a stick to beat Hayes with and is laying it on too hard.

The third question is most difficult of all and guarantees subjectivity on the commentator's part. So, except for a small technical comment, I will be subjective completely and openly. I believe it is a shame for two of us to enter into fiery disputes when there are so few of us who do long range *comparisons(4) and* fewer still who care about and believe in the Austric hypothesis.*(5)* Many of us have been involved in squabbles about theory or method or first authorship. Everyone knows that I have been in more than my share. Yet these are not desirable relationships for long rangers to have with each other. Even from past issues of the [MT] Newsletter or the Journal, we know that the Blust-Hayes squabble has been going on for some time. With no inside knowlege of either person's psyche, we still can tell from the behaviors that there are key issues. On the one hand Blust and his colleague Lawrence Reid seem to deny Hayes' existence or scholarly contribution. That is enough to anger anyone; it can outrage others. On the other hand, Hayes appears to have reacted so fiercely that both Blust and Reid probably got angry themselves and continued to ignore Hayes. Words have flown forth, and maybe back, and have struck soft psyches at tender points. The words get public, and reputations are damaged or threatened.

What to do? Well, I invoke the spirit of Paul Benedict, who would have both of them say ... "Anywaaay, let's forget that and get on with what we agree about!" We do not need to fight with each other. There are plenty of idiot linguists out there to fight with. Find a good pub. Sit down together and have a tasty pint of good Czech lager! And relax! (Bomhard tells me that I might even be able to do that with Trask!)

The technical comment was that Hayes would have been helped a lot had he had access to Blust's new massive reconstruction of proto-*Austronesian(6) Hayes* now has access to it and he appears to be willing to modify his argument somewhat in accordance with the new information. I reckon that Paul Benedict, who dearly loved a spirited friendly argument, would have decided that Hayes had made a good point, nay a powerful one and a very useful one. Of course I cannot prove that.

---

4. In a recent manuscript submitted for publication elsewhere, Merritt Ruhlen has proposed dropping the distinction between `long range' and `short range' comparisons. Since he believes that the methodology is the same in either case, he fears that using `long range' implies a new method or unorthodox or unsatisfactory methods. Not caring overly much about methodological theory, I still use the different terms because they differ considerably in the amount of work and difficulty involved and because short range relationships are usually pretty evident or obvious.

5. I mean `believe' in the sense that someone thinks that a given hypothesis is probably true, considering the empirical basis for it and the fruitfulness of it. I do not mean `believe' in a religious or philosophical sense, a matter taken on faith or as part of a world view.

6. In our SPECIAL ISSUE of *Mother Tongue* (October 1999) we mentioned Blust's paper on Austronesian, which was basically an announcement of the massive work on proto-Austronesian. Jared Diamond recently had a piece on Austronesian in *Nature* (vol. 403, February 2000: 709-710), saluting Blust's work and discussing Pacific prehistory, which now had important new conclusions because of Blust's work. Surely his arguments on Malayo-Polynesian sailing and especially Polynesian sailing in outrigger `canoes' is a beautiful contribution to Oceanic and to global prehistory. My thanks to Ofer Bar-Yosef and his staff for apprising me of Diamond's article.

Response to Blažek's Comments

By La Vaughn H. Hayes

Dr. Blažek's comments provide an excellent illustration of some of the points I try to make in my response to Blust about the intricacies and difficulties of the Austric comparison, but he seems to have missed some significant points of my Austric sibilants paper. I refer specifically to his conclusions that the arguments for AA */z/ and its correspondence to AN */D/ are unconvincing and that he can accept only the "hair," "wind/rain," "bark of tree/root," "root," "water," and "tail" comparisons. I propose in the paper systematic correspondence of both AA */s/ and */z/ to AN */h/ and */D/, respectively, and the "hair" comparison evidences in fact */s/, not */z/. The "root" comparison should be excluded, as noted in my response to Blust, its inclusion being due to error on the writer's part and the editor's dissemination of the draft to commentators before that error could be corrected. But that is not all.

Blažek also observes that without careful AA reconstruction based on daughter-language reconstruction, AA/AN comparisons remain only on the speculative and uncertain level. This statement is true to a certain degree, but we are still years, if not decades, away from the day when bottom-up reconstruction in Austroasiatic will have reached the point that such careful AA reconstruction is possible. In the interim, there is no theoretical reason why top-down reconstruction, as essayed in the writer's Austric series and the sibilants paper, cannot proceed or produce useful and scientific results which are neither speculative nor uncertain.

As stated in Austric I (Hayes 1992:148-9), a crucial problem of the Austric comparison is to determine whether partial correspondence due to common origin exists between Austroasiatic and Austronesian. Such genetic correspondence includes recurrent correspondence of phonemes, and a key aim of the sibilants paper is to show that such correspondence exists between the AA sibilants, */s, z/, and AN */h, D/, respectively. To demonstrate this existence, the lexical data is presented in groups of comparisons, each based on a specific type of regular and recurrent AA/AN correspondence involving specific reflexes of the Austric sibilants.

The patterns of phonological correspondence demonstrated by that arrangement and their implications should be clear to all, but curiously, both Blažek and Blust either do not perceive them or they simply ignore them. Blust deals with the matter in one way (see elsewhere in this volume); Blažek resorts to the old method, often seen in literature dealing with discussion of megalo-comparisons, of introducing additional lexical data which purportedly shows why a view opposite to the writer's is justified. This method has the advantage of introducing new data and stimulating thought and/or conversation; its disadvantage is that objectivity is all too often replaced by subjectivity, for the assumption is that in absence of demonstration of *bona fide* genetic correspondence, one opinion is as good as any other. That is not the case here.

Take the "night/dark" comparison, for example, which Blažek finds unconvincing. The writer routinely limits his exemplary data citation to three glosses in order to keep the paper's size small, hence cites only Proto-Waic *som 'night', Khasi *kajiŋ dum* 'night', and VN *ʔdəm* 'dark', the first to evidentiate a reflex of AA *zəm, the latter two of *nzəm. Possibly, Blažek has interpreted this to mean that those are the only available examples, but that is hardly the case. The writer has 15 additional examples from 12 other languages or dialects which support ***zəm** and 11 additional examples from 8 other languages which support ***nzəm**, cf. Bateg Nong *ləsəm*, Che' Wong *səm* 'rain', Khmu' (Yuan) *pəsəam, pasəam*, Lawa *saom* 'night', Mon *kəsɐm* 'rainy season', Pacoh *lsam* 'winter', Riang (White Striped) *khiin sɑm*, Riang (Black) *təm khiin shɑm* 'midnight', Theng *pəsuom* 'night', Vietnamese *toy tɐm* 'be very dark', *thəm* 'be black, black and blue, (lips) blue', *thəm* '(of color) be dark', Wa (Tung Va) *grəŋ sɑm* 'midnight', Santali *bəsyəm* 'a.m./morning' and Jeh *see rədaam* 'darkening', Khmu' *dɨm* 'stay overnight', Mon (Old) *døm* 'lodge for the night', Palaung *dɯm* 'to lodge', Vietnamese *ʔdɐm, ʔdem* 'dark', Kharia *nuḍum* 'early before dawn', Mundari *nudum* 'twilight', Santali *aḍam jaḍam* 'throughout the night, every night', Santali *hadam hudum* 'dusk, dawn', *kaḍam kaḍam* 'moving about at night, roam about at night'.

The evidence supporting reconstruction of AA ***nzəm** is thus relatively massive and (I think) entirely persuasive, but Blažek argues that Proto-Waic ***som** is instead comparable with Mon samò, Proto-South Bahnaric ***mhʌ:** 'evening' and Palaung *hmȫ* 'night, evening'. These forms may indeed be comparable, but if so, one must propose ***zəmo** or the like as their antecedent, an affixed variant of the ***zəm** that has undergone a different pattern of stress shift, whence ***s(ə)mo** > Mon *samò,* etc. While plausible, Blažek's replacement comparison must be regarded as even more speculative than the writer's, and still no real just cause to reject the latter.

Similar comments could be made about the other comparisons Blažek addresses in his comments, but I see no point in doing so. The basic issue is the question whether or not the lexical and phonological data presented by the writer adequately supports his hypothesis about the evolution of the Austric sibilants and its implications for the existence of Austric. I think the answer must be in the affirmative because the regular and recurrent correspondence shown in the paper to exist between the AA/AN reflexes of the Austric sibilants cannot be ignored or simply explained away by new comparisons that may or may not exhibit such correspondence. The existence of such genetic phonological correspondence is not just an anomalous bizarrerie of nature; it is an icon of diachronic linguistic change, and, if it has no meaning, then historico-comparative linguistics has no meaning and we are all wasting our time.

Response to Blust's Comments

By LaVaughn H. Hayes

Dr. Blust is or should be well known for his vitriolic attacks on the work of long-rangers. His criticism of the late Paul K. Benedict's work on Austro-Tai is only slightly less harsh than the present critique of the writer's Austric sibilants article. While Blust undoubtedly means well, he seems determined to impose unrealistically high standards on these comparisons. This predilection is understandable. Austronesianists like Blust have had an easy time of it, because Dempwolff handed them a full-blown AN [Austronesian] reconstruction on a platter. They have never had to face the groundbreaking problems, difficulties, and uncertainties of a new long-range comparison, because Dempwolff also took care of all that for them and it only took him 30 years or so to do it. With the luxury of a mature, well-researched comparison as one's baseline, it is apparently easy to become a high priest of methodological dogma and technical precision and indulge oneself in fastidious evaluations, while failing to realize that the harsh criticism is apt to do more harm than good to the targets of one's scorn.

Blust's discourse on logic and technical proofs is interesting and illuminating, but it amounts to a bit of overkill in the present context, as some of his other comments do also. A word should be taken at face value, i.e. in its more common sense, if it is not otherwise qualified. As the writer's dictionary defines it, proof is merely the evidence establishing the validity of a given assertion, and it is in that common, non-technical sense that he uses the word 'proof' in the Austric sibilants article. But of course Blust needs proof for his own assertion that the article is not science and consequently must first interpret proof in a different and specific way.

Otherwise, the writer is led by the content and nature of Blust's comments to conclude that Blust has probably never read any of the writer's other publications concerning Austric and spent far less time on the article under discussion than he should have.

Despite the lecture, Blust abandons all logic in order to reach the conclusion that "The present work is not a piece of science." More than a very casual reading of the article should make it clear to any observer that all of Blust's charges leading to that conclusion are false. It is untrue that the comparative method has not been used and equally untrue that the exhibited comparisons are not based on recurrent phonological correspondence. How indeed can one even make such an accusation as the latter, when all of the comparisons presented in the body of the article are grouped on the very basis of specific regular and recurrent phonemic correspondences, such as AA */z/ to AN */D/?

Blust's next conclusion that "What we see is a grasping at straws, not evidence of recurrent sound correspondences" is even less understandable, but one begins to perceive the roots of his misunderstanding in the statements supporting that conclusion. Blust is apparently unaware that the writer has published around 400 AA/AN lexical comparisons, all of which exhibit regular and recurrent sound change and correspondence to some degree. This quantity constitutes only about a third of the writer's Austric comparative database, so the old idea that lexical evidence does not exist to support the Austric unity can be abandoned. One problem with these comparisons, which the writer has addressed in his articles, is that many of the lexical similarities can indeed be explained as products of borrowing or chance. The importance of the Austric sibilants article resides in the very fact that it presents similarities that cannot plausibly be explained away in that fashion, but Blust has evidently failed to grasp that most significant fact.

Blust stands on firmer ground when he criticizes the writer's usage of AN source materials. The writer's primary source for AN proto-form citation is Wurm and Wilson 1983, which is a multi-source compilation of PAN and lower level AN reconstructions. For a variety of reasons, the writer is unable to evaluate the quality of these reconstructions, hence has been more or less forced to pick and choose the proto-forms which seem most useful for his purposes. Blust focusses on Dyen's revision of Dempwolff's work, but Dyen's reconstructions, which are cited in Wurm and Wilson, are confusing to the non-specialist because different versions of the same proto-form are given from different Dyen publications. The writer has been unable to obtain copies of Dyen's work and learn what the variation means, and, for those reasons, he has tended to avoid Dyen's proto-forms.

Over the years, the writer has become aware of a number of problems with AN reconstruction, some of which are probably unknown even to the Austronesianists, but dealing with such problems is really beyond the scope of his research, the primary purposes of which are to discover AA evidence of the Austric unity and the diachronic evolution of Austroasiatic. As for Blust's own work, the writer uses Blust's published reconstructions when pertinent, but he does not have access to the ACD and has never seen any notice about it on the Internet nor heard anything about how to access it from any Austronesianist he has been in contact with.

The writer might add in passing that over the past 20 years, he has sought on a number of occasions to open a dialogue with Austronesianists, including Blust, in order to get needed assistance with AN reconstruction and other questions, but has met with little willingness to cooperate. This is most unfortunate, for the the Austric comparison has some important things to say about AN reconstruction, which is not quite as perfect as Blust and others seem to think.

Blust is correct that PAN *biRaq should be used vice *biRah 'alocasia (species)'. However, inclusion of this comparison is the writer's error, and it should be deleted. In any case, no Austric sibilant is involved.

Otherwise, Blust's criticism of the writer's comparisons probably reveals more about his own lack of understanding of the intricacies of AA linguistic history and the difficulties of the Austric comparison than it does about the infelicities of the writer's comparative methodology. Compared to Austroasiatic, Austronesian has apparently been conservative in its diachronic evolution, thus making AN reconstruction a comparatively easy exercise. In Austroasiatic, things are quite different, for massive change has been the rule and attrition of initial and final elements of the lexical word the norm. Mysterious patterns of stress shift have apparently operated, moving back and forth in largely unfathomable fashion, also causing medial elements to drop out. Given such circumstances, it is hardly surprising that monosyllabic AA forms may correlate with one AN syllable in one comparison and with another in another comparison. The AA correspondents of the other AN syllables in any given comparison are usually absent because of diachronic attrition, but there is nothing the comparatist can do about it.

As a consequence of the nature of AA diachronic evolution and the general absence of higher level AA reconstruction, the Austric comparison is admittedly a risky business subject to frequent error, but accusing the comparatist of irresponsibility or a lack of integrity for trying to make the best of a bad situation is more overkill and prevents not a single error. If the phonology correlates and the semantics are on target or close, it is reasonable to propose lexical correspondence. Try that with unrelated languages and see how far you get. But that procedure is in essence no different from what Austronesianists or other comparatists do. The difference lies in the fact that they usually have more pieces of the puzzle to work with, thus are less susceptible to commission of error.

In light of the inherent difficulties and problems of studying Austric, the writer must conclude that Blust's comments will not do much to help the writer in his work and in fact serve little useful purpose at all. This is unfortunate, for what is needed for the future is mutually beneficial cooperation between Austroasianists and Austronesianists, not contempt or scorn by one side for the other's efforts. One can take it on faith or one can take it on the massive amount of linguistic evidence the writer has presented publicly, but Austroasiatic and Austronesian are genetically related. This fact will become clearer over time, but the sibilant correspondence confirms the proposition now in a unique and apparently irrefutable way, and the writer feels that more open-minded and attentive observers will find themselves able to agree with that conclusion.

Response to Fleming's Comments

By LaVaughn H. Hayes

Dr. Fleming's comments are on the whole the sort of reasoned, seasoned, and eminently fair critique I have come to expect from him over the years on any topic, and I find little to object to or rebut in them. However, I am disappointed that he sides with Robert Blust in alleging that I seem "not to understand what science really does." Blust's allegations are a canard, and I would prefer to ignore them. But unfortunately they exist and have already won one convert, so there is no escape: They must be acknowledged and answered in the interest of damage control.

There seems to be no easy answer and certainly no simple one to such pernicious allegations. The fundamental issues are philosophical, not merely scientific or linguistic, as we shall see below. I shall endeavor, nonetheless, to present an adequate response in minimal space.

First, let me say, for the edification of Messieurs Fleming and Blust, that I was introduced to what science does and how it does so (the scientific method) in high school nearly 50 years ago. I had cause to relearn both lessons in college, where I earned a bachelor's degree in mathematics and computer science, and often to review them during the 17-year course of my Austric research. I doubt that either Blust or Fleming knew all those facts, and certainly neither one asked me about my background before leaping to their negative conclusions in quite unscientific fashion.

As one can see from Blust's comments on my Austric sibilants paper, those conclusions are based primarily on one small thing, Blust's objection to my usage of the word "proof". To justify his objection, Blust goes to pains to delineate his concept of science, and how scientific research is conducted. He fails to mention, however, that his concept is only one of many, and he does not inquire or even speculate about the possibility that the writer uses one different from his own. In brief, Blust is convinced that he has the truth, and the only truth, on his side.

As the writer learned it, science seeks to construct an accurate representation of the world, and its means of doing so is called the scientific method. This method entails four steps: observation, hypothesis, prediction, and testing. In producing the sibilants paper, I followed all of those steps: (1) I observed certain phenomena (the various correspondences presented in the paper), (2) constructed a hypothesis to explain what certain phenomena (the sibilant correspondences) mean, (3) used the hypothesis to predict other phenomena, namely the evolution of the Austric sibilants and the genetic relatedness of Austroasiatic (AA) and Austronesian (AN), and (4) tested the hypothesis to see if it accounted for the phenomena accurately and adequately. Publication of the paper in this journal makes it available to other scholars, who may test my findings, if they so desire.

While I was testing the Austric sibilant hypothesis, the only plausible explanation I could find for the existence of the unique correspondence of the AA and AN reflexes of the Austric sibilants is inheritance from a common ancestor. Since I was unable to find any evidence that contradicts that explanation, I had to conclude that the indicated genetic relatedness could not be disproved on the basis of the evidence available to me, which is why I called that unique correspondence "irrefutable proof" of the existence of Austric. That conclusion does not exclude the possibility that counterevidence will not be found in the future, in which case the explanation would have to be modified, but until such evidence is found, the explanation can be accepted as scientifically correct. In fact, I hoped that someone else might be able to refute or further verify my theory about the sibilant correspondence evidentiating the Austric relationship, one important reason for publishing the paper, but thus far, the comments on the paper have been focussed elsewhere.

Now, unless I am badly mistaken, many, if not most, observers will find absolutely nothing wrong with my application of the scientific method, nor cause therein to doubt that I know reasonably well what science does, so why the big fuss and all the shabby allegations by Blust? The reason is simply this: he espouses a different vision of science and scientific method, as implied above. That vision is based on the philosophy of Karl Popper, author of *The Logic of Scientific Discovery;* and an enormous difference in concept and world view is involved.

The Popperian scientific method also has four steps, which differ from those described above, but they need not concern us here. What is of interest is the underlying precepts. Popper believes that there are no scientific statements which are not open to question, because absolute truth lies outside the bounds of scientific knowledge. Hence, he advocates exactly what Blust contends, that there is no such thing as scientific proof, and that is evidently why Blust makes such a big issue of my usage of the word "proof." It is also why he leaps to all his other sad conclusions about my lack of knowledge of science and my work not being science.

Popper also believes that disprovability or refutation is actually the fundamental source of scientific knowledge. Thus, scientific statements must be tentative and disprovable. If you see the sun shining above, you can never conclude that it is shining above unless you can prove that it is not shining above. This point of view seems so wrongheaded and contrary to reality to this writer that he feels that he surely must be missing something important and necessary to a proper understanding of Popper's philosophy. But it does further explain Blust's violent reaction to my advocacy of irrefutable proof, though one should note that Blust never bothers to disprove my proof directly, and hence does not practice his own scientific method as it should be practiced. Should I now allege that he does not understand what science does and deem all his work unscientific?

There are other differences between us and also an apparent paradox. Popper says that a theory is scientific if, and only if, it is capable of being refuted. Thus, if Blust has truly refuted my theory that the sibilant correspondence proves the existence of Austric, then he has also verified that it actually is scientific. In short, he has shown his allegation that the paper is not science to be false, an outcome of using the Popperian method which the writer finds both bizarre and hilarious.

In conclusion, this matter is less a question of who is right and who is wrong than it is about differences in the philosophy being used. The crux of the conflict is not whether I or Blust know what science does and how to do it (clearly, we both do), but our respective concepts of what scientific truth is. Blust and the Popperians do not believe that man can know the real world with certainty; they condemn him to an eternal guessing game. If the sun is shining above, that is not a real truth for them, only a high number on the probability scale. The writer and many others take an opposite view that man can know reality on the basis of logical and material facts obtained through empirical observation and reason. Those facts are the evidence which provides the proof for any assertion about the real world.

A Comparison of Basque and (North) Caucasian Basic Vocabulary

John D. Bengtson
156 15th Avenue NE
Minneapolis, MN 55413
U.S.A.
john.bengtson@co.hennepin.mn.us

These Basque-(North) Caucasian lexical comparisons are based on the Basque words in the expanded (200 + 19) Swadesh list provided by R.L. Trask (1997, pp. 352-357). The Caucasian forms (with a few exceptions) are drawn from the Caucasian dictionary (NCED) by Nikolayev & Starostin (1994), and the book on West Caucasian (CWC) by Chirikba (1996). PNC stands for the deepest Caucasian reconstruction (Proto-North Caucasian), PEC (Proto-East Caucasian) and PWC (Proto-West Caucasian) stand for the two lower level subdivisions of that family. Citation of reconstructions should not be taken as my approval or endorsement of technical details of the reconstructions, but only as hypotheses of the respective authors.

In the lexical comparisons that follow, the Basque word is cited first, followed by the symbol (=), meaning "is compared with," and then by Caucasian words and reconstructions that *may* be cognate. In a few cases, alternative comparanda are offered.

1. 'back': Basque *bizkar* (*bi-*zka*-r) = PAbkhaz ***zakʷa** > Abkhaz a-*zkʷa* 'back (with shoulder blades)' (CWC 200) [Burushaski *-sqa* 'on one's back'] § For Basque initial *bi*-and final *-r,* see Appendix C: Fossilized Morphological Elements.

2. 'bark': Basque *azal* 'skin, bark' = Abaza *čʷa* 'skin, bark', Budukh *ʕič* 'skin (of cattle)', etc. < PNC ***ʔwā̆rćwə** 'skin, color' (NCED 228) § Comparison requires the analysis of Basque *azal* as *aza-l,* assuming *–l* to be a suffix, as in the next comparison.

3. 'belly': Basque *sabel* 'belly, stomach' = Bezhta *šebo* 'liver', Chechen *žim* 'kidney', etc. < PNC ***ʒ̆ăwV** 'kidney, gall' (NCED 1106) § For Basque *–l,* cf. the previous comparison.

4. 'bird': Basque *txori* [čori] ~ *xori* [šori] = Chamalal *č̣or* ~ *č̣oru* 'bird', Avar *č̣orólo* 'quail', etc. < PEC ***č̣HwīlV** 'a small bird' (NCED 388)

5. 'bone': Basque *hezur* ~ (Z) *ézür* ~ (R) *ẽzur* 'bone' = Rutul *sur* 'part, side' (<'*rib'), Archi *bars:on* 'rib', Lak *niws* 'rib', etc. < PEC ***rɨ̄mswe** (implicitly ~ ***mswɨ̄re**) 'side, rib' (NCED 954)

6. 'breast': Basque *bular ~ bulhar* 'breast, chest, bosom; mother's milk' = Avar *ʁʷári* 'udder', Hunzib *ʁeru* 'breast, udder', Lak *q:ʷar ~ q:ʷal* 'udder', etc. < PEC ***ɢwălħē** (NCED 465) § For Basque *–r,* see Appendix C.

7. 'claw': Basque *aztapar > atzapar* = Dargwa *ṭuṗ* 'finger', Lezgi *ṭub* 'finger', *tuṗal* 'finger-ring', etc. < PEC ***ṭwɨbi** 'finger' (NCED 1007) § Analyzed as *a-z-tapa-r,* with initial and final fossilized morphological elements (see Appendix C). Only the segment *–tapa-* is compared with PEC ***ṭwɨbi.**

8. 'cold': Basque *hotz* = Hinukh *=oč̣č̣u,* Hunzib *=ɔč̣č̣u,* Adyge *čəʔa* 'cold', etc. < PNC ***č̣wErHV / *rHEč̣wV** 'cold' (NCED 393)

9. 'come': Basque *jaugin ~ jin* = Avar (Chadakolob) *=áʕn-* 'to come', Andi *=uɢVn-* 'to enter (pl.), to become', Hinukh *=aq̇-* 'to come', etc. < PNC ***huq̇Ŭn** 'to go, to come' (NCED 611)

10. 'day': Basque *egun (~ egur- ~ egu-)* = Lak *q̇ini,* Archi *iq,* Khinalug *q̇ä* 'day', etc.< PEC ***Hwīq̣V** 'day' (NCED 622) [Burushaski *gon* 'dawn', *gunc* 'day']

11. 'die': Basque *hil* = Chechen *=al-,* Karata *=ilʔ-,* Adyge *ƛ̣a-n,* etc. 'die' < PNC ***=iwƛ̣Ĕ** 'to die, to kill' (NCED 661)

12. 'dirty': Basque *zikin* = Kryz *č̣äq* 'eye secretion', Ubykh *c̣aq̇ʷ* 'manure, dung', etc. < PNC ***č̮Hɨq̇wĂ** 'dung, ordure, dirt' (NCED 387)

13. 'dog': Basque *or* ~ (Z) *hor ~ ho* = Kabardian *ħa,* Avar *hoj,* Budukh *χor,* Dargwa (Chirag) *χ̮:wa,* etc. 'dog' < PNC ***χHwěje** (NCED 1073) § In PEC, the oblique stem ended in *–r-,* thus Budukh *χor* (above), Godoberi *χʷar-di* 'dogs', etc.

14. 'dry': Basque *agor (~ igar ~ eihar)* = Avar *=aq̇ʷ:ara-b,* Hunzib *qoqoru,* Udi *q:ari,* etc. 'dry' < PNC ***=iɢ̱wĂr** 'dry, to dry' (NCED 631)

15. 'ear': Basque *belarri (be-larri)* = Chechen *lerg* (dial. ~ *lerig ~ larig*), Dargwa **leħi > liħi ~ laħi,* etc., Abaza *ləmħa* etc. 'ear' < PNC ***łĕHłe ~ *łĕHłi** (NCED 756) § For Basque *be-,* see Appendix C.

16. 'earth': Basque *lur (lurr-)* = Avar *raƛ̣:* 'earth, ground', Lak *luχč:i* 'earth, land', etc. < PNC ***lhĕmŁ̱wɨ̆** 'earth' (NCED 747) § For assimilation or dissimilation of *l ~ r,* see Appendix B.

17. 'eye': Basque *begi* = Chechen *bʕärg,* Avar *ber,* Udi *pul,* Dargwa *ħuli,* Ubykh *bLa,* etc. 'eye' < PNC ***ʔwĭlʔi** 'eye' (NCED 250) § See Appendix C, and Bengtson 1999 for detailed phonological derivation of Basque *begi*.

18. 'fall' (v) (1): Basque *erori* = Lak *t:iri=x:i-* 'to fall, scatter', Tabasaran *a=x-* 'to fall' (< PLezgian ***ˀarλɨ-**), etc. < PNC ***HraλwE** 'to fall, go down' (NCED 602)

19. 'fall' (v) (2): Basque *jausi* = Chechen *=oss-* 'to descend, go down', Hunzib *=õs-* 'to fall down, be scattered', etc. < PEC ***=ŭ̄śV** 'to descend, fall, be scattered' (NCED 1011)

20. 'far': Basque *urrun* ~ (Z) *hürrün* = Rutul *χɨri*-dɨ, Lak *arχ-,* Agul w-*arχa,* etc. 'far' < PEC ***=ārχV** 'far' (NCED 269)

21. 'fat/grease': Basque *ziho* = Tindi *c̣inλu-,* Dargwa *c̣erx:-* 'fat' (adj) < PEC ***c̣ēnx̱wV** ~ ***c̣ēnλwV** 'fat' (adj) (NCED 362)

22. 'father': Basque *aita* = PWC ***(a)tʷə** 'father' > Ubykh *tʷə,* etc. (CWC 390)

23. 'few': Basque *guti* ~ *gutti* ~ *gutxi.* [guči] = Lezgi *güṭü* 'narrow', Budukh *goda* 'short', Dargwa (Kaitag) *kuṭil* ~ (Chirag) *kuṭze* 'short', etc. < PNC ***ḵHə̄ṭwV̄ / *ḵwHə̄ṭV̄** (NCED 690)

24. 'fire': Basque *su* [śu] = Lak *c̣u,* Bezhta *c̣o,* Khinalug *č̣ä*, Abaza *m-ca,* etc. 'fire' < PNC ***c̣ăjɨ̆** 'fire' (NCED 354) [Burushaski *ši* 'fireplace, hearth']

25. 'fly' (v): Basque *hegaz* egin (egin = 'do') = Tsakhur al-*iχas,* Dargwa (Chirag) *iX̰-,* etc. 'to fly' < PEC ***HiχV** (NCED 582)

26. 'fog': Basque *laino* ~ *lanbro* = Chechen *doχk* 'fog', Avar *naḳ* 'cloud', Khinalug *unḳ* 'cloud', etc. < PEC ***rănƛ̱wV̆** 'cloud, fog' (NCED 947)

27. 'foot': Basque *oin* ~ (Z) *hun* [huñ] 'foot' = Chechen *ħaqolg* 'ankle(-bone)', Lak *niq:a* 'heel', etc. < PEC ***ʔĭnɢwV̆** (NCED 248) [Burushaski *-γan* 'heel']

28. 'four': Basque *laur* ~ *lau* = PWC ***p(:)əƛ̣́ə** 'four'; cf. PEC ***bū̆nḺe** 'eight' (NCED 314) [Burushaski *alt-* '2', w-*ált* '4', *alt*-amb- '8']

29. 'freeze': Basque *izoztu* (*izoz*-tu) = Abaza *cħa-šʷə* 'ice', Khwarshi *=ujc̣-* 'cold', etc. < PNC ***=əc̣̱ŏr-** 'to freeze' (NCED 419)

30. 'go' (walk): Basque *joan (i-oan)* = Karata *=oˀan-* 'to go', Batsbi *ʕo-* 'to go' (pres.), Hurrian *un-* 'to come', etc. < PEC ***=VˀwVn** 'to go' (NCED 1016)

31. 'guts': Basque *hertze* ~ *heste* ~ (Z) *erxe* [érše] = Avar *ʕorčo- / ʕarča-* 'guts' § Word apparently restricted to Avar. Comparison by René Lafon (1952).

32. 'hair' (1): Basque *ile* ~ *ilhe* ~ (B) *ule* 'hair' ~ (Z) *ilhe* 'wool' = Tsakhur *arX̰* 'autumn wool', Batsbi *ajχ*-nĭ 'woollen thread, yarn', etc. < PEC ***ʔālχV** 'wool' (NCED 242)

33. 'hair' (2): Basque *bilo* ~ *bilho* '(single) hair' (Z distinguishes *bilho* 'human hair' from *ilhe* 'wool') = Lak *ṗiħulli* 'feather', Abaza *bra* 'mane', Kabardian *baLā-ca* 'shaggy', etc. < PNC ***ṗVħVɬV** 'feather, mane' (NCED 879) [Burushaski *pholγó* 'feather']

34. 'hand': Basque *esku* = Khinalug *čigin* 'shoulder', Tsakhur *guč* 'arm', Khwarshi *geša* 'arm, foreleg', etc. < PEC ***gwəćV** (implicitly ~ ***ćəgwV** > Khinalug *čigin*) 'arm' (NCED 448)

35. 'head': Basque *buru* = Udi *bul* 'head', etc. (see NCED 1041) § The PEC etymology ***wĕnƛ̣V** seems to reflect a merger or contamination of two words that remain distinct in Basque (*buru* 'head' ~ *mutur* 'snout') and Burushaski (*bur* '(single) hair' ~ *-multur* 'nostril').

36. 'hear': Basque *entzun* = Batsbi *=abc̣-* 'to know, get to know', Dargwa *=umc̣-* 'to search', Ubykh *ć̣a-* 'to know', etc. < PNC ***=ămc̣Ĕ** 'to know, see' (NCED 262) § For 'hear ~ know', cf. 'know' (a fact), below, no. 45.

37. 'heavy': Basque *astun* = Chechen *stomma* 'thick, dense', Lak *c:un-* 'dense', etc. < PEC ***c̱ŭjmV** 'thick, dense' (NCED 328)

38. 'here': Basque *hemen* ~ (BN, Z) *heben* = probably a combination of demonstrative elements corresponding to PEC ***hă** [NCED 486] + PNC ***mV** [NCED 842]; cf. Tabasaran *hamu* 'this', Tsez *he-me*-si 'that' (far, invisible), etc.

39. 'hit': Basque *jo (i-o)* = Chamalal *=uʁ* 'to hit', Tabasaran *u=ʁ̮* 'to break', etc. < PNC ***=HiʁwE(r)** 'to beat' (NCED 569)

40. 'horn': Basque *adar (< *ardar)* = Avar *ƛ:ar* 'horn', Chechen *kur* 'horn', etc. < PEC ***ƛ̱wɨ̆rV** (NCED 771) [Burushaski *tur* ~ *-ltur* 'horn']

41. 'how' (1): Basque *zela* ~ *zelan* (based on interrogative stem *ze-*) = Hunzib *su* 'how!', *su* di 'anyhow', PCircassian ***sə-tʰə** 'how?', etc. < PNC ***s̱āj** 'what' (see 'what', below)

42. 'how' (2): Basque *nola* (based on interrogative stem *no-*) = Hunzib *hi-na* 'how?', etc. < PNC ***hīnV** oblique interrogative stem (see 'when', below)

43. 'I': Basque *ni* = Lak *na,* Dargwa *nu* 'I' < PEC ***nɨ̆** (NCED 855) § In Caucasian, this word remains only in Lak and Dargwa: see the discussion in NCED.

44. 'knee': Basque *belaun* ~ *belhaun* ~ (Z) [bélhañ] 'knee', [belháyko] 'on one's knees' < **be-lhaur-iko* = Akhwakh *eƛelo* 'elbow', Tsez *horu* 'elbow', Agul q:ar-*xil* 'elbow', etc. < PEC ***ƛ̆wilV** 'elbow' (NCED 770) § For Basque *be-,* see Appendix C.

45. 'know' (a fact): Basque *jakin (i-akin)* = Akhwakh *=eq̇-* 'to know', Khwarshi *=iq̇-* 'to know', Dargwa *=aq̇- / =iq̇-* 'to hear', etc. < PNC ***=ĭq̇Ē** 'to know, hear' (NCED 646) [Burushaski *hákin, -ki-* 'to learn']

46. 'know' (a person): Basque *ezagu*-tu = Avar *c̣eχ:é-* 'to search, ask', Ubykh *ʒʁa-* 'to ask', etc. < PNC ***c̣EnχV(n)** 'to search, ask' (NCED 359)

47. 'leg' (1): Basque *zango* ~ (Z) *zankho* < **s̄an-ko* = Lak *s:an* 'foreleg, paw, pad', (Khosrekh) *s:ana* 'wedge', Rutul *sɨn* 'front part of leg', etc. < PEC ***s̱ɨ̄nŏ** 'long bone' (NCED 963) [Burushaski *sesen* ~ *susun* ‘elbow’: cf. Udi *sun* ‘elbow’]

48. 'leg' (2): Basque *hanka* ~ (Z) *anka* = Tindi *anq:u* 'knee-bone', Andi *aq:u* 'thigh', Archi *aq* 'hind leg of an animal; leg, foot', etc. < PEC ***ʔānq̱V** (NCED 244)

49. 'male': Basque *ar (arr-)* = Ingush *ärh* 'ungelt', Lak b-*urx*-ni 'male', etc. < PEC ***ˀīrƛ̱wV** 'male' (NCED 210) [Burushaski *hir* 'man, husband']

50. 'man': Basque *gizon / giza-,* Aquitanian CISON (a name) = PAbkhaz ***qaḉ̣a** 'man' (CWC 389)

51. 'many': Basque *asko* = Andi *-eč̣uχa* 'big', Lak *č̣a̰u* 'many', Ubykh *šχʷa* 'strong', etc. < PNC ***č̣HəqwV** 'big' (NCED 386) [Burushaski *c̣ik* ~ *c̣iq* 'all, altogether']

52. 'moon': Basque *ilargi (il-argi)* ~ (Z) *argizagi (argi-zagi)* as to the element *argi* = Lezgi *raʁ* 'sun', Lak *barʁ,* Khinalug *ɨnq̇,* Kabardian *dəʁa,* etc. 'sun' < PNC ***wirə̄q̇Ă** (NCED 1051)

53. 'narrow' (1): Basque *estu* = Dargwa *č̣a̰rṭa,* Andi *č̣:iṭir* 'narrow', etc. < PEC ***č̣̰HVrdV** (NCED 387)

54. 'narrow' (2): Basque (AN, G) *medar ~ mehar ~ mear* = Hunzib *ƛamu* 'thin, narrow', Chechen *nilχa* 'thin, sparse', etc. < PEC ***ɦimV̄ƛ̠̣V̆*** (NCED 521)

55. 'near': Basque *hurbil* 'near' (adj) = Chechen *gerga* 'near', Avar *ʕagára*-b 'near', etc. < PNC ***ɦĭgVrV*** 'near' (NCED 518)

56. 'night': Basque *gau, gab-* = Tsakhur *χ̮am* 'night', Tabasaran *χ̮ab*-aq̰ 'evening', Agul *χ̮aw*-aq̰ 'evening', etc. < Proto-Lezgian ***χ̮:am:*** (Chirikba 1985: 103, adding also Abkhaz [Bzyb] *-xa* < **-qa* 'night' [?])

57. 'not': Basque *ez* ~ (B) *ze* = Chechen *ca* 'not', Karata *-č̣e* (negative particle), Rutul *ǯ*- (general negative), etc. < PEC ***ʒ́ə̆*** / ***ć̣ə̆*** 'not' (NCED 1101)

58. 'old': Basque *zahar* = Lezgi *sur* 'old', *jis* 'year', Chechen *šira* 'old', *šo / šera-* 'year', Ubykh *žʷə* 'old', etc. < PNC ***swĕrho*** 'old, year' (NCED 968)

59. 'root': Basque (c) *erro* ~ (Z) *herro* 'root', also 'teat' (of udder), 'ray' (of sun) = Avar *rix* 'vein, blood vessel', Lak *x:ʷa* 'sinew, tendon, string', etc. < PNC ***x̱wɨ̄ˀrV*** (NCED 1064)

60. 'round': Basque *biribil* ~ *borobil* = Avar *gʷangʷára* 'skull', Hunzib *gogor* 'cheek', Agul *gurga*-ħ 'skull', etc. < PEC ***gwɨ[l]gwə*** 'round object; skull' (NCED 450)

61. 'sand': Basque *hondar (hond-ar)* = Khinalug *ant* 'earth, ground', Tsez *atu* 'dirt, mud' < PEC ***ˀantV*** (NCED 201)

62. 'see': Basque *ikusi* = Bezhta *=ĩq-* 'to find', Budukh *irq-* 'to see', etc. < PEC ***=Hārq̇V(n)-*** 'to see, find' (NCED 547)

63. 'sew': Basque *josi (i-osi)* = Andi *=eš:-* 'to weave', Dargwa (Urakhi) *=irš- / =uš-* 'to weave', Adyge *śa-n* 'to weave', etc. < PNC ***=irś̠E(l)*** (NCED 653)

64. 'sharp': Basque *zorrotz* = PAbkhaz ***ć̣arə*** 'sharp' (CWC 390; cf. NCED 1045-1046)

65. 'sister': Basque *ahizpa* 'sister (of woman)' < **a-hiz-pa* = Bezhta *is* 'brother', *isi* 'sister', Tindi w-*ac:i* 'brother', j-*ac:i* 'sister', etc. < PNC ***=ĭć̠ĭ*** 'brother, sister' (NCED 669) [Burushaski *-c̣u* 'brother, sister'; as to the Basque element *-pa / -ba* (cf. also arre-*ba* 'sister [of man]', al(h)a-*ba* 'daughter', etc.), cf. PAbkhaz ***pa*** 'son' > Ashkharywa a-*pá,* Tapant *pa,* etc.]

66. 'sit': Basque *jarri (i-arri)* = Tsakhur ǵ-*i-ʔar-* 'to sit, sit down', Abkhaz *a-ja-rá* 'to lie', etc. < PNC ***=eʔ(w)Vr** 'to sit, be quiet' (NCED 409)

67. 'skin' (1): Basque *larru* (~ B *narru*) 'skin, hide, leather' = Avar *ƛ̣:er* 'color', Dargwa (Akushi) *guli* 'skin, sheepskin', etc. < PNC ***Łŏli** 'skin > color > paint' (NCED 789)

68. 'skin' (2): Basque *azal* 'skin, bark': see 'bark', above.

69. 'sky': Basque *zeru* ~ (R) *zeuru* ~ *zéuri* ~ (Z) *zélü* 'sky' = Avar *zob* 'sky', Lak *s:aw* 'sky', *zʷal* 'god', Chechen *dēla* 'god', etc. < PNC ***ʒə̆wɨ̆** 'sky', PEC ***ʒwǟlV** 'god' (NCED 1092, 1097)

70. 'sleep': Basque *lo* (egin) = Akhwakh *ƛ̣:unu-* 'to sleep', Khwarshi *ƛe*-s- 'to sleep', mu-*ƛu* 'dream', etc. < PNC ***=HVƛ̣̱wĀn** 'to sleep' (NCED 619)

71. 'small': Basque *tipi* = Avar *hiṭína*-b 'small', Archi *ṭi* 'small', Tsez *ṭaḳi* 'few', etc. < PNC ***ṭiHV / *HiṭV** 'small, little' (NCED 1001)

72. 'snow': Basque *elur* ~ (Z) *elhur* [élhür] = Chechen *lō,* Batsbi *law,* Lezgi *žiw,* etc. < PEC ***jĭwλV / *λĭwV** 'snow' (NCED 684)

73. 'spit' (v): Basque *listu* (egin) = Tindi *λac:i* 'saliva', Tsez *λač̣:i* 'pus', etc. < PEC ***ƛ̆ămVč̣V** (NCED 767)

74. 'squeeze': Basque *hertsi* = Rutul *=ir(i)č̣a-* 'to press, squeeze', Dargwa (Chirag) *=ač̣-* / =*ạ̰č̣*- 'to press, squeeze', etc. < PNC ***Hič̣Ăn** (NCED 568)

75. 'star': Basque *izar (izarr-)* = Tindi *c:aru,* Bezhta *cã,* Dargwa (Chirag) *zure,* etc. < PNC ***ʒwhărī / *ʒwăhrī** 'star' (NCED 1098)

76. 'stick': Basque *makila* 'stick, staff' ~ (B) *maket* 'stout stick, club' = Bezhta *maq* 'stake', Andi *moq̇:ol* 'ceiling', Ingush *beq̇a* 'pole', etc. < PNC ***bħə̄nq̱̇V̆** 'pole, post' (NCED 295)

77. 'stone': Basque *harri* = Akhwakh *χaχi* 'road metal', Lezgi *χirχem* 'small stones, road metal', etc. < PEC ***χHĕrχV** 'small stone, gravel' (NCED 1073) [Burushaski *xóro* 'small stones']

78. 'sun': Basque *eguzki (egu-*zki) ~ (Z) *ékhi* = Archi *iq* 'day', etc. (NCED 622: see 'day', above)

79. 'tail': Basque *buztan* = Tsakhur *bɨ̰t* 'tail' < PLezg ***p:o̰ç̣:-** (NCED database) § Isolated? cf. Yeniseian **pis* '(bird's) tail'.

80. 'that': Basque *hori* (mesial), *hura* (distal) 'that' = Andi *hu*-du- 'that', Karata *ho*- 'that', etc. < PNC ***ʔu** ~ ***hu** (demonstrative pronoun) 'that' (NCED 222)

81. 'there' (mesial): Basque *hor* = Budukh *wo-rə*-n 'there' (above speaker), Kryz *wa-rä* 'there' (above speaker), etc. < PNC ***ʔu** ~ ***hu** (see 'that', above)

82. 'there' (distal): Basque *han* = Dargwa *hanna* 'now', etc. < PNC ***h[ä]nV** (NCED 487; probably connected with PEC ***hă** : see NCED 486, and 'this', below)

83. 'this': Basque *haur* ~ *hau* = Chechen *ha-ra* 'this', Botlikh *ha-* 'this', etc. < PEC ***hă** emphatic demonstrative stem (NCED 486)

84. 'thou' (1): Basque *hi, -k* 'thou' (intimate) = Chechen *ħo,* Dargwa (Akushi) *ħu* ~ (Kaitag) *i,* Udi (Nidzh) *hu*-n 'thou', etc. < PEC ***ʁwV̄** 'thou' (oblique base) (NCED 483) [Burushaski *go-* 'thine', *gu-/gú-/gó-* 'thou']

85. 'thou' (2): Basque *zu* 'thou' (unmarked: historically derived from a second person plural pronoun, cf. English *you,* French *vous,* etc.) = Chechen *šu* 'you' (pl.), Lak *zu* 'you' (pl.), etc. < PNC ***źwĕ** 'you' (2nd pers. pl.) (NCED 1086)

86. 'tie': Basque *lotu* = Lezgi *iliṭ*-iz 'to bind around', Kryz *ju-ṭul-* 'to tie, bind', etc. < PNC ***jeṭal-** 'to tie, bind; untie' (NCED 679: implicitly ~ ***jelaṭ-** > Lezgi)

87. 'tongue': Basque *mihi* ~ *mii* ~ (B) *min* (< **mihi*-n) = Tindi *mic:i,* Rutul *miz / mizi-,* etc. < PNC ***mĕlç̱ĭ** 'tongue' (NCED 802) [Burushaski *-mélc̣* 'jaw'] [See Bengtson 1999 for a complete phonological derivation of Basque *mihi*]

88. 'tooth' (incisor): Basque *hortz* = Lak *k:arč:i* 'tooth', Akhwakh *goržo* 'fang, canine tooth', etc. < PEC ***gəlǯwē** 'fang, canine tooth' (NCED 435-436: the reconstruction ***-l-** seems uncertain: "We must reconstruct some medial liquid ... most probably ***-l-**," though this is circumstantial, since all languages with liquids preserved have *-r-*. The Basque word, with trilled *-r-*, would fit well with PEC/PNC ***kərǯwē > *gərǯwē,** with assimilation.)

89. 'tooth' (molar): Basque *hagin* = Bezhta *hagna* ~ *χäg* 'molar tooth', Hunzib *χɔʁin* 'molar tooth', etc. < PEC ***qămkV** (NCED 883)

90. 'tree': Basque *zuhaitz* ~ (Z) *zuhaintze* (compound of *zur* 'wood' + *hai[n]tz[e]*):

Basque *zur* ~ (B) *zul* 'wood' = Avar *ç̣ul* 'wood, firewood', Botlikh *ç̣uli* 'stick', etc. < PEC ***ç̣wɨ̄ɬɦV̆*** 'stick, branch' (NCED 374)
Basque *-haitz* ~ *-haintze* 'tree' = Hinukh *aže* 'tree', Karata *eže*-la 'pine tree', Adyge *čə-ɣə* 'tree', etc. < PNC ***Hă(r)ǯwī*** 'a kind of tree' (NCED 549)

91. 'turn': Basque *itzuli* = Agul (Burshag) *ilcan-* 'to turn (on an axis)', Chechen *χ-arc-* 'to turn, overturn (intr.), etc. < PEC ***=īrcVl*** 'to twirl, turn round' (NCED 649)

92. 'twenty': Basque *hogei* = Avar *q̇:ó*-go, Lak *q̇u,* etc. < PEC ***ɢ̱ə̆*** 'twenty' (NCED 456)

93. 'two': Basque *bi* ~ *bi-ga* = Udi *p:a̰,* Ubykh *ṭq̇ʷa,* Khinalug *ḳu,* etc. < PNC ***q̇Hwā̄*** 'two' (NCED 924); Chirikba postulates PWC ***dɢʷə*** (CWC 395)

94. 'warm': Basque *bero* 'hot' = Tsez *boboru*, Khwarshi *bobolu* 'hot' (Bokarev 1959: 241); and cf. PWC ***bla/ə*** 'to burn' (CWC 393)

95. 'water': Basque *ur* ~ (Z) *hur* = Lezgi *hül* 'sea, liquid', Avar *ʕor* 'river', etc. < PEC ***ɦwɨ̆lV*** 'river, reservoir' (NCED 537); or, alternatively: Lezgi *wir* ~ *ür* ~ *hür* 'lake, pond', Avar *ħor* 'lake, pond', etc. < PEC ***ħwirɨ*** (NCED 506) [Burushaski *hur* 'wooden water conduit']

96. 'what': Basque *zer* (interrogative stem *ze-)* = Ingush *se* 'what', Avar *s:un-* 'what' (oblique), Ubykh *sa* 'what', etc. < PNC ***s̱āj*** 'interrogative pronoun (what)' (NCED 958) [Burushaski *-sa* ~ *-se* (in *bésa* ~ *bése* 'why': cf. Archi *basa* 'when']

97. 'when': Basque *noiz* (interrogative stem *no-)* = Tsez *neti* 'when', *na* 'where', Lezgi *ni* 'who' (erg.), etc. < PNC ***hīnV*** 'oblique interrogative stem' (NCED 491-493)

98. 'where': Basque *non* = Tabasaran *naʔan* 'where', Hunzib *nijo,* Tsez *na* 'where', etc. (see 'when', above) [Burushaski *áne* 'where']

99. 'who': Basque *nor* = Lezgi *ni* 'who' (erg.), Agul *hina* ~ *na* 'who' (erg.), etc. (see hen', above)

100. 'wing': Basque *hegal* [heɣal ~ eɣal] = Lak *qa* 'wing', Tabasaran *χil* 'hand', etc. < PEC ***q̱ɨ̆lʔi*** 'elbow, arm, wing' (NCED 895); or, alternatively: Agul (Tpig) *χil* 'wing', etc. < PEC ***χĕɬHe*** 'sleeve' (NCED 1070)

101. 'woman' (1): Basque *emakume* (ema-*kume*) = Lak *qami* 'women', Khinalug *χin-* 'woman', etc. < PNC ***qwănV*** (NCED 900) [*eme-* in both Basque words for 'woman' means 'female']

102. 'woman' (2): Basque (Z) *emazte* (ema-*zte)* = Chechen *stē* 'woman, female', *stē*-n 'female', *steš* ~ *stij* 'women, females', Ingush *se* 'woman, female', *istij* 'women, wives'; Dargwa (Chirag) *cade* 'female' < PEC ***cVjdV** ~ ***ćVjdV** (NCED 375); or, alternatively: Chechen *zuda* 'woman', Urartian *ašti* 'woman, wife', etc. < PEC ***ʒHV̆ṭV** (NCED 1094)

103. 'woods' (1): Basque *oihan* = Chechen *ħun* 'forest', Avar *xon*-λí 'shady side, slope', Tindi *han*-da 'hill', etc. < PNC ***fănV** 'mountain, hill' (NCED 425) [Burushaski *hun* 'wood, firewood']

104. 'woods' (2): Basque *baso* 'woods', *basa-* 'wild' = Akhwakh *beča* 'mountain', Tindi *besa* 'mountain', Archi *sob* 'mountain pasture', etc. < PEC ***wɨ̄ce** (NCED 1053)

105. 'worm': Basque *har* ~ (R) *ãr* = Avar *ħapára* 'worm', Tsakhur *a̰bra*-wučẹ 'worm', etc. < PEC ***ɦabarV** 'worm' (NCED 508)

106. 'ye': Basque *zuek* (2nd person plural pronoun: plural of *zu*) = Lak *zu* 'you' (pl.), etc. (see 'thou' [2], above)

Analysis and Conclusions

The above corpus consists of 106 Basque-Caucasian comparisons. There is of course some overlap (i.e., two or more Basque words are compared with the same Caucasian morpheme, e.g., 'day = sun', 'how [1] = what', 'how [2] = when, where, who', etc.). In 56 cases, Basque words are matched with words that are (according to Nikolayev & Starostin and/or Chirikba) common to the whole (North) Caucasian family (PNC); in the remaining 50 comparisons, Basque words are compared with words restricted to a lower-level subdivision: East Caucasian (PEC) or West Caucasian (PWC); and in a few cases ('back, guts, I, man, night, tail') the Basque word is matched with words restricted to one or two Caucasian languages. Note, however, that the majority of comparisons are Basque-PNC.

The comparisons can also be analyzed according to degree of semantic exactness. A comparison is judged semantically exact if the Basque meaning is the same as the meaning of the Caucasian reconstruction *or* at least one of the attested languages cited. In semantically inexact comparisons the meanings are slightly different, but historically relatable, e.g. 'hand ~ arm', 'bone ~ rib', 'knee ~ elbow', etc. Of the 106 comparisons, 78 (74%) are semantically exact, and 28 (26%) are semantically inexact. Note again that exact comparisons predominate, indeed make up about ¾ of the total. If we restrict the count to the 56 Basque-PNC comparisons, the proportion of semantic exactness is somewhat greater: 45 exact (80%) and 11 inexact (20%).

In terms of the 219 meanings in Trask's list, Basque and Caucasian matches total 48% (106 of 219, including semantically inexact and taxonomically isolated comparisons). Restricted to the most tightly constrained set of comparisons (the 45 that are semantically exact and common to Basque and PNC), the percentage of matches is 21%. According to Swadesh (1954), 12-36% of matches indicates a stock of about 25 to 30 centuries of divergence, and a relationship that should be "obvious to the linguist." This would be a chronological level comparable in depth to Proto-Indo-European or Proto-Sino-Tibetan, a level Peiros (1997) calls "Old Family."

Let me emphasize that I am not a mathematician nor a lexicostatistician, and the interpretation given in the preceding paragraph is meant only as a very rough and impressionistic approximation. Probably the strongest statement I am entitled to make is that my interpretation of the above body of comparisons indicates that Basque and the (North) Caucasian language have a significant amount of basic vocabulary in common, and that this evidence points in the direction of a genetic relationship, particularly when one takes into account the recurrent phonological correspondences and apparent fossilized morphological elements outlined below (see Appendices A, B, C). I will leave more precise calculations of genetic affinity to my colleagues who are more qualified to do this kind of work.

Appendix A: Phonological Patterns

Phonological regularity, in itself, does not "prove" genetic affinity, but regularity of correspondence in *basic* words is an indicator of real relationship as opposed to random resemblance. Within this set of comparisons it is possible to find many cases of recurrent phonological correspondences between Basque and Caucasian, for example:

1. Basque trilled *rr* [r̄] (usually written –*r* in final position in the standard orthography: e.g. *izar* [is̄ar̄] 'star', *izarra* 'the star') corresponds to PNC ***r**, e.g. PNC ***ʒwăhrī** 'star'; see also 'bone, dry, far, horn, male, old, root, sharp, sit, stone, worm'.
2. Basque tapped *r* (written as single *r* between vowels, also some final *r*'s, as in *(h)ur* 'water', *zur* 'wood') corresponds frequently to PNC ***l** or ***ł**: see 'bird, knee, round, sky, tree, warm, water'.
3. PNC ***Ł**, ***ƛ**, ***ƛ̄** (lateral affricates, roughly = [dl, tl, tl']) correspond to Basque *l,* in initial or final position: see 'die, fog, four, skin (1), sleep, spit'.
4. The internal Basque cluster –*lh*-, preserved only in northern Basque (Basse Navarre, Lapurdi, Zuberoa), has two main counterparts in Caucasian:
    - (a) PNC cluster of ***l** /***ł** + laryngeal or postvelar fricative, e.g.:
      Northern Basque *bulha-r* = PEC ***ɢwălħē** (see 'breast')
      Northern Basque *ilhe* = PEC ***ʔālχV** (see 'hair' [1])
      Northern Basque *bilho* = PNC ***ṗVħVłV** (see 'hair' [2])
    - (b) PNC ***λ** (voiceless lateral fricative):
      Northern Basque *elhur* = PEC ***λĭwV** (see 'snow') There is only one example of (b) in this set of etymologies, but cf. also: Northern Basque *ilhinti* 'firebrand' = PEC ***λwindV** '(fire)wood' (NCED 764); Northern Basque *olho* 'oats' = PNC ***λwɨ̆ʔwV** 'millet' (NCED 763), etc.
5. Basque *b* corresponds to PNC (PEC/PWC) ***w**, as well as some clusters containing ***w** (***gw**, ***ɢw**): see 'breast, eye, fog, head, round, two, woods (2)'.
6. Basque *z* (lamino-alveolar [s̄]) corresponds clearly to PNC ***s**: see 'bone, how (1), leg (1), old, what'; and to PNC ***ʒ**: see 'sky, star'.

Appendix B: Irregular Changes:
Metathesis, Assimilation, Dissimilation

The acceptance of some of the comparisons above relies on the recognition that seemingly "irregular" processes are also at work in linguistic change, namely *metathesis, assimilation*, and *dissimilation.*

Metathesis is quite frequent in Caucasian languages, apparently because their phonetic systems are so complex, that the relative *order* of phonetic segments is less important than in languages with simple systems. A single Caucasian language or dialect can have two variants of the same common word: for example, in the Akushi dialect of Dargwa, 'tongue' can be *mez* or *lezmi,* while in most other Dargwa dialects one or the other variant is preferred: Chirag *mec:,* Urakhi *miʒ,* vs. Kadar *limzi,* Kaitag *luc:umi,*etc. (see NCED 802). The phonetic system of present-day Basque is much simpler than that of any Caucasian language, but even in Basque some metathetic variants are known between the dialects. 'Liver' is *gibel* in most Basque, but also *bigel* in Alto Navarro. It should not be surprising that if there are cognates between Basque and Caucasian, some would require metathesis, thus:

Basque *(h)ezur* 'bone' vs. PEC ***rɨ̄mswe*** 'side, rib' (but Rutul *sur* presupposing
PEC ***mswɨ̄re***);
Basque *laino* ~ *lanbro* 'fog' vs. PEC ***rā̄nƛ̱wV̆*** (*lanbro* would fit well with
PEC ***ƛ̱ā̄nwrV̆*** !);
Basque *bilho* 'hair' vs. PNC ***ṗVħVłV*** 'feather, mane' (but Burushaski *pholγó*
'feather', with same order as Basque);
Basque *esku* 'hand' vs. PEC ***gwəćV*** 'arm' (but Khinalug *čigin* < ***ćəgwV***)
Basque *zahar* 'old' vs. PNC ***swĕrhŏ*** 'old, year';
Basque *lotu* 'to tie' vs. PNC ***jeṭal-*** (but Lezgi *iliṭ-* presupposes ***jelaṭ-***).

Assimilation and *dissimilation* come into play mainly in words that have two liquids (or lateral affricates) in both languages:

Basque *lur (lurr-)* 'earth' vs. PNC ***lhĕmḺwɨ̆*** > Avar *raƛ̣:*;
Basque *erori* 'to fall' vs. PNC ***HraƛwE*** 'to fall';
Basque *be-*lhaur-* 'knee' vs. PEC ***ƛwilV*** 'elbow';
Basque *larru* (~ Bizkaian *narru*) 'skin,' etc. vs. PNC ***Łŏli*** > Avar *ƛ̣:er.*

In these examples we see the interchange of laterals and rhotics, as well as alternation of ordinary laterals (l) with lateral fricatives (λ) or lateral affricates (ƛ, L). Without external comparison, it is difficult or impossible to tell which language (if any) retains the original form. In the last example we find another kind of dissimilation *(l>n)* in Bizkaian Basque.

In two cases in the above sample we find another kind of dissimilation: when the original cluster **-rd-* in Basque is followed by *r,* the first *r* drops out. So Basque *adar (adarr-)* 'horn' < **ardar,* and *medar* 'narrow' < **merdar.* Elsewhere (e.g., Bengtson 1995:91-92, 1997:144-145) I have shown that the Caucasian lateral affricates (ƛ,ƛ,L) correspond regularly to Basque *l-* in initial position, but to Basque *–rd-* in medial position. (The clearest example of the latter is probably PNC ***=ĕƛ̣Ĕ*** 'middle, half' = Basque *erdi* 'middle, half'.) In the case of Basque *medar* ~ *mehar* > *mear* 'narrow', the latter two forms are apparently due to contamination with Basque *mehe (> mee)* 'thin'.

Appendix C: Fossilized Morphological Elements

Throughout this corpus of lexical comparisons we find some apparent petrified or fossilized morphological elements, some of which have been noted for decades by C.C. Uhlenbeck, Karl Bouda, and others. These fossilized elements were discovered when it was noticed that there were recurrent "leftover" segments – both initially and finally – in Basque words compared with their Caucasian counterparts. The elements are of two main types: (a) fossilized case prefixes (Basque *be-/bi-*, *e-/i-*, *a-*, etc.); and (b) fossilized plural ending (Basque *-r*).

Apparent fossilized prefixes in Basque words have been noticed by several scholars, notably C.C. Uhlenbeck (1927), who wrote about Basque *be/bi-*. Examples from the above body of comparisons are:

Basque *bi-zka-r* 'back' = PAbkhaz ***zakʷa** 'back', Burushaski *–sqa* 'on one's back'
Basque *be-larri* 'ear' = PNakh ***-lari** (oblique stem) 'ear', etc.
Basque *be-gi* 'eye' = Chechen *bʕärg* 'eye', etc.
Basque *be-lhaun* ~ *be-lhaur-* 'knee' = PEC ***λ̄wilV** 'elbow'

Other probable fossilized prefixes include Basque *e-/i-* (see *e-gun* 'day', *e-sku* 'hand', *e-lhur* 'snow', *i-zar* 'star') and Basque *a-* (see *a-dar* 'horn', *a-hiz-pa* 'sister'). Other examples of the Basque fossilized prefixes, their proposed correlations with the Proto-Caucasian system of class markers, and arguments against the criticisms of Trask (1995) are found in my earlier articles (Bengtson 1995, 1997).

Another "leftover" element is the final *-r* found in some Basque words, but not in the proposed cognate Caucasian stems. For example, from the comparisons above:

Basque *bi-zka-r* 'back' = PAbkhaz ***zakʷa** 'back', Burushaski *–sqa* 'on one's back', etc.
Basque *bul(h)a-r* 'breast' = PEC ***ɢwălħē**
Basque *honda-r* 'sand' = Khinalug *ant*, etc.
Basque *e-l(h)u-r* 'snow' = Chechen *lō*, etc.

other examples:

Basque *nega-r* ~ *niga-r* 'tears' = Udi *neʁ* 'tear', etc. < PEC ***nĕwqŭ** (NCED 848)
Basque *ziga-r* 'mite' = Dargwa (Akushi) *ç̣ika* / *ç̣eka* 'flea', etc. (NCED 376)
Basque *bela-r* 'forehead' = Rutul *bäl* 'forehead', etc. (NCED 285)

fossilized *–r* in Basque can plausibly be identified with the Caucasian plural ending ***-r.** This ***-r** persists as a plural marker in some Caucasian languages, for example Hunzib:

Hunzib *koma-r* 'kidneys', *qõqla-r* 'eggs', *ɑṗɑ-r* 'paws', *xõṭox-ar* 'frogs', *har-ur* 'eyes', *ca-b-ur* 'stars', etc. (van den Berg 1995)

Plural ***-r** also remains in certain sporadic fossilized forms in Caucasian languages, plural in form but singular in meaning, for example:

Tabasaran *marc-ar* 'hearth (NCED 308)
Khinalug *cul-oz* 'tooth' (*-oz* < ***-or**: NCED 326)
Dargwa (Akushi, Chirag) *nerʁ* 'tear' < ***neʁʷ-r** (NCED 848)

The last example is entirely parallel with Basque *nega-r* ~ *niga-r* 'tear(s)' (see above), except that Dargwa has taken the further step of metathesizing the last two consonants *(ʁr > rʁ).*

References

Bengtson, J.D. 1995. "Basque: an orphan forever? A response to Trask." *Mother Tongue* I:84-103.

---------- 1997. "Basque and the Other Dene-Caucasic Languages." In *The Twenty-Third LACUS Forum,* ed. by A.K. Melby, pp. 137-148. Chapel Hill, NC: LACUS.

---------- 1999. "'Eye, Ear, Tongue' in Basque and East Caucasian." In *From Neanderthal to Easter Island (Tribute to W.W. Schuhmacher),* ed. by N.A. Kirk and P.J. Sidwell, pp. 3-10. Melbourne: Association for the History of Language.

Berg, Helma van den. 1995. *A Grammar of Hunzib.* Munich/Newcastle: LINCOM EUROPA.

Bokarev, E.A. 1959. *Tsezskie (didojskie) jazyki dagestana.* Moscow: Akademija Nauk.

Chirikba, V.A. 1985. "Baskskij i severokavkazskie jazyki." In *Drevnjaja Anatolija,* pp. 95-105. Moscow: Nauka.

---------- 1996. *Common West Caucasian: The reconstruction of its phonological system and parts of its lexicon and morphology.* Leiden: Research School CNWS.

Lafon, René. 1952. "Etudes basques et caucasiques." *Acta Salmanticensia: Filosofia y Letras* 5/2: 5-91.

Nikolayev, S.L., and S.A. Starostin. 1994. *A North Caucasian Etymological Dictionary.* Moscow: Asterisk.

Peiros, Ilia. 1997. "Macro-Families: Can a mistake be detected?" In *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin,* ed. by I Hegedűs, P.A. Michalove, and A.M. Ramer, pp. 265-292. Washington, DC: Institute for the Study of Man.

Swadesh, Morris. 1954. "Perspectives and Problems of Amerindian Comparative Linguistics." *Word* 10: 306-332.

Trask, R.L. 1995. "Basque and Dene-Caucasian: A critique from the Basque side." *Mother Tongue* 1: 3-82.

---------- 1997. *The History of Basque.* London/New York: Routledge.

Uhlenbeck, C.C. 1927. "Die mit *b-* anlautenden Körperteilnamen des Baskischen." In *Festschrift Meinhof.* Hamburg.

Unusual Phonetic Characters

ã, etc.	nasalized vowels
a̰, etc.	pharyngealized vowels = NCED aI, etc.
c:, q:, s:, etc.	tense obstruents = c̄, q̄, s̄, etc.
c̣	(Caucasian) lax glottalized hissing affricate = c’
c̣	(Burushaski) retroflex hushing affricate
č̣	lax glottalized hushing affricate
ć	voiceless hissing-hushing affricate
ć̣	lax glottalized hissing-hushing affricate
ɢ	voiced uvular stop
H	laryngeal of undetermined quality (in reconstructions)
ħ	voiceless emphatic laryngeal fricative
ɦ	voiced laryngeal fricative
ʔ	glottal stop
ʡ	glottalized emphatic laryngeal stop
ʕ	voiced emphatic laryngeal fricative
ḳ, ṗ, q̇, etc.	glottalized stops = k’, p’, q’, etc.
χ	voiceless uvular fricative = x̣
χ̰	pharyngealized voiceless uvular fricative
ł	lateral resonant or glide (in reconstructions)
λ	voiceless lateral fricative = ɬ
ƛ	voiceless lateral affricate = tɬ
ƛ	lax glottalized lateral affricate = tɬ’
Ł	voiced lateral affricate = dl
ʁ	voiced uvular fricative = γ̣
ʁ̰	pharyngealized voiced uvular fricative
ś, ź	hissing-hushing fricatives
ʒ	voiced hissing affricate = dz
ʒ́	voiced hissing-hushing affricate
ǯ	voiced hushing affricate = ǰ

Abbreviations

AN	Alto Navarro (Basque dialect)
B	Bizkaian (Basque dialect)
CWC	*Common West Caucasian* (Chirikba 1996)
G	Gipuzkoan (Basque dialect)
NCED	*North Caucasian Etymological Dictionary* (Nikolayev & Starostin 1994)

PEC	Proto-East Caucasian
PNC	Proto-(North) Caucasian
PWC	Proto-West Caucasian
R	Roncalés (Basque dialect)
Z	Zuberoan = Souletin (Basque dialect)

Some remarks on „A Comparison of Basque and (North) Caucasian Basic Vocabulary“ by John D. Bengtson

Václav Blažek
Ryneček 148
26101 Příbram III
Czech Republic
e-mail: blazek@phil.muni.cz

Comments on specific comparisons:

bone The reconstruction **enazur* ~ **anezur* of Michelena based on R ẽ*zur* (cf. DELV XIII: 165) is in perfect agreement with PECc **mswɨre*.
fog In DELV XIX: 35-36 there are these forms: V, G *laiño* "niebla", AN, G "paño de agua que cubre los cristales", L *lainho* "enfermedad del trigo", AN, L, BN *lan(h)o* "niebla, bruma", S *lanhũ*, L *laino* id. The other forms, e.g. *lanbro* "bruma, llovizna" (> Gasc *l'embrun* more probably than vice versa), *lanbera* "aguanieve, agua que cae de las nubes mezclada con agua", L *lanbo* "llama, bruma, oscuridad, nube", S *lanpu* "niebla" (DELV XIX: 42), besides V *landur(a)* "llovizna", AN *lankar* "id., lluvia menuda", V *lantar* id. (DELV XIX: 44), appear to be compounds of **lan/ñ-* with various components.
freeze The verb is derived from *izotz* N, V, L, R, S "escarcha", R "rocío", L, R "agua de la niebla". Michelena (see DELV XVIII: 96) analyzes it as a compound consisting of *ihi(n)tz* "rocío" & *(h)otz* "frío". I think he is right. On the other hand, the forms L *ihintz*, AN, G, R *intz*, BN, S *ihitz*, V *euntz*, *iruntz*, *iñuntz* "rocío" (**iNitz* ?) are compatible with PNCc **jansē* "hoar-frost, snow" or PNCc **jǎmʒA* "snow" (NCED 675, 674). It was Trombetti (1926: 131) who already compared the Basque word with Andi *anži* "snow" < **jǎmʒA*. (Why was this comparison not included?)
gut The Avar counterpart is too isolated.
hair$_1$ DELV XV: 194 quotes N *ille* "pelo", N "lana", S *ilhe* "lana", L "ciertos pelos cuya raíz penetra en la garganta del cerdo y puede ahogarlo (enfermedad),“ AN *elle*, Aezc. *eile* "lana, pelo", V *ul(l)e*, (RS) *hule* "pelo". Especially the latter forms resemble Gothic *wulla* "wool" (after Uhlenbeck).
hear The reconstruction *entzun* "oír", V, BN "sentir, percibir olores", V "famoso, célebre" < **e-nezu-n* of Michelena (DELV X: 40) seems to be compatible with PNCc **-ǎmc̣E*. On the other hand, the closeness of Fr *entendre* "to hear", past participle *entendu,* cited by H. Schuchardt looks very suggestive.
know Schuchardt (RIEV 7: 336 & 10: 158 - see DELV XII: 147) connected Basque *ezagutu* with Gasc *sa(u)gut*, Prov *sa(u)but* < Vulgar Latin **saputu* from *sapio,-ere* "wissen" (cf. M-L # 7586).
leg$_2$ There are two sets of Romance parallels which can represent a source of the Basque word: (1) Cors, Sard *anka* "Bein" < Latin *ancus* "gekrümmt" (M-L #446); (2) Sp, Pg, Ct, Prov, It *anca*, Fr *hanche* "Hüfte" (M-L #4032) < Low German *hanke* "hind leg of a horse", cf. English *haunch* < (O)Fr.

many In DELV IV: 60, 65 there is *asko* derived from *asi* "hartarse" plus suffix *-ko*. Semantically it is fully legitimate, cf. Czech *dosti* "enough" derived from *syt* "satiated, replete".

narrow$_2$ There are important related forms in Basque: BN, L, S *hertsi*, AN *ertsi*, *ertxi* "estrecho, angosto" (DELV XI: 75), better compatible with PECc **č̱HVrdV*. Bengtson offers another solution for *hertsi,* comparing it with PNCc **Hič̣An* "to press, squeeze". (Why was the *-r-* of Rutul *irič̣a* not projected in the reconstruction [by Nikolayev & Starostin]?)

sand In DELV XXI: 188 there is a rich spectrum of meanings of Basque *(h)ondar*, viz. "residuo, sedimento-heces; V, G "arena, playa, borra de líquido", AN, L, BN, R "último", AN, L, BN "fondo". Especially the latter meaning gives support for derivation from Sp *hondo* < Latin *fundus*, *-i* /*-oris* (cf. M-L # 3585).

stick Rijk (Lingua 12: 232) proposed a connection of Basque *makil(l)a* with Hebrew *maqqēl* id. Could it indicate Phoenician / Punic influence ?

tail Following Corominas (see DELV VII: 91), it is necessary to take in account such forms as Sal *uzterina*, S *üztari*, *buztarin* "grupa" (Couvr. "cola") which can really be derived from Vulgar Latin **pŏstĕrio, -ōne* > OFr *poistron*, OIt *postione* "Hintere" (M-L #6688). I see an alternative etymology in the separation of the *b*-prefix plus **(u)ztan* "back", cf. *izter* "thigh", *iztai* "groin" (so Lafon quoted in DELV), Tibetan **s-tiŋ* "behind, after" or sim.

woods Add Basque *oyana* "monte" [closer to Caucasian meanings such as „hill, shady slope“] (DELV XXI: 179).

worm Berger (MSS 9[1956]: 7) compared Basque *har* "Made" with Burushaski *hər* "Kornwurm". On the other hand, there are dialectal forms such as R *anr* = *ãr, ã(n)ra* "gusano de la carne, col" leading to Michelena's reconstruction **(h)anar*, while Corominas reconstructs **anăra* , quoting also Bigorra *ario* (< *árĕa* < **árana*), supposing the same origin for Ct *arna*, Prov *arno* "Motte, Milbe" (DELV III = ASJU XXII-3: 895; M-L # 8586). The forms with nasals support Bengtson's comparison with PECc **ɦabarV*, maybe as follows **Habr(a)* > **HaMr(a)* > *hã̆r(a)*.

Conclusions:

If the parallels are to demonstrate a close genetic relationship, I am ready to accept only the safe cognates. That is why some meanings are rather expressive for this demonstration (e.g. the semantic variability in the case of "dirty", further "hit", "round", "spit", "turn" etc.). I think that the following etyma belong among the safe, probable or at least hopeful comparisons:

bird bone breast cold come day die dog dry ear earth eye fall$_{1\ \&\ 2}$ **far fat few fire fly fog foot four hair**$_2$ **hand head horn I knee know ?male moon narrow/squeeze night not old root sew sister sit skin sky ?sleep** (if it is not a Lallwort) **snow star stone sun that this thou**$_{1\ \&\ 2}$ **tongue tree ?twenty** (too high a numeral is always suspicious) **two warm water what when where white who wing woman**$_{1\ \&\ 2}$ **woods**$_{1\ \&\ 2}$ **worm ye**

Altogether 66-68 parallels. I find most significant the large number of grammatical words (pronouns, adverbs), and the class prefixes.

The comparison of a modern language (Basque) and a reconstructed protolanguage (Proto-North Caucasian) is rather dangerous for several reasons: c e.g. the semantic reconstructions can be only approximative. Taking in account only two living languages, say Basque and Lak, and 100-word list with exact semantic correspondences, there are 8 common cognates: **fire I star thou tongue two what woman**. This result is comparable with results for the comparison between, e.g., Indo-European and Uralic within the Nostratic macrofamily, cf. e.g. Finnish and German: *kuka // was* "who", *mies (*mäńćä) // Mann* "man", *nimi // Name* "name", *paljon // viel* "many", *puu // Baum* "tree", *tämä // diese* "this", *tuo // der* "that", *vesi // Wasser* "water", cf. also *minä* "I" *// mich*, *mir* "me".

Abbreviations

Basque dialects (after Azkue): Aezc Aezcoano, AN alto navarro, BN bajo navarro, G guipuzcoano, L labortano, R roncalés, S suletino, V vizcaíno; Cc Caucasian, Ct Catalan, E East, Fr French, Gasc Gasconian, It Italian, N North, O Old, P proto-, Prov Provençal, Sp Spanish

References

AASJU *Anejos del Anuario del Seminario de Filología Vasca "Julio de Urquijo".*

DELV *Diccionario etimológico vasco* by Manuel Agud & Antonio Tovar. Donostia - San Sebastian 1989f.

M-L Meyer-Lübke, W. 1935: *Romanisches etymologisches Wörterbuch.* Heidelberg.

MSS *Münchener Studien zur Sprachwissenschaft.*

NCED *A Northern Caucasian Etymological Dictionary* by S.L. Nikolayev & S.A. Starostin. Moscow: Asterisk 1994.

RIEV *Revista Internacional de Estudios Vascos.*

“Vasconic” and Basque-Caucasian

By W. Wilfried Schuhmacher
Kirkebakken 13
4621 Gadstrup
Denmark

Theo Vennemann’s recent articles (see last but not least Vennemann 1998) posit three language families in prehistoric Europe north of the Alps:

Old European
Atlantic (i.e., Afro-Asiatic)
Indo-European

The Old Europeans speaking “Vasconic” (related to contemporary Basque) moved, about 10,000 BP, from southern France into the rest of the continent (also eastward), where their language(s) became adstrata or substrata.

As for possible Basque (*recte:* Vasconic) - Caucasian comparisons, therefore, a “differential diagnosis” has to be made:

1. Are the parallels due to a genetic relationship dating from before10,000 BP?
2. Or are the parallels due to adstratum/substratum influence from after 10,000 BP?

(One hypothesis would naturally exclude the other.) Diagnosis (1) would get more weight if the parallels are supported by other Dene-Caucasian families/languages, and diagnosis (2) would be relevant if the parallels are *not* attested by evidence from families/languages outside the Basque-Caucasian area.

How complicated the matter might turn out to be might be seen from the case of ‘apple’ – irrelevant in this connection – as the IE word might be a loan from Semitic or – Burushaski ...[1]

Reference

Vennemann gen. Nierfeld, Theo. 1998. “Basken, Semiten, Indogermanen: Urheimatfrage in linguistischer und anthropologischer Sicht.” In: *Sprache und Kultur der Indogermanen,* ed. by Wolfgang Meid, pp. 119-138. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

---

[1] [Editorial note] What Schuhmacher refers to here is English *apple,* Old Irish *aball,* Oscan *Abella* (malifera), Lithuanian *óbuolas*, Polish *jabłko,* etc., which have been compared with Burushaski *balt* ‘apple’ (< *baƛ ?). Greek μᾶλον ~ μῆλον > Latin *mālum,* Italian *mela,* etc., have a nasal labial instead. (Cf., metathesized, Turkish *elma,* Hungarian *alma,* etc.)

Assessing Basque-Caucasian and Other Remote Relations: an increasingly cautious view.

Paul Sidwell

Australian National University

I was interested to receive John Bengtson's letter of August 23rd 1999, inviting me to comment on his claim that one can find 29% ~ 41% cognacy between Basque and Caucasian basic vocabulary. I perused the attached list of lexical comparisons, and, despite the many striking resemblances that have been compiled, I still find myself agreeing with Trask (1995 and passim) that there is "zero evidence" for a genetic relation between these widely separated languages. The following is my 'devil's advocate' answer to that letter, and hopefully it will only be one event in a lively and helpful dialogue.

Perhaps it will disappoint some of my colleagues, but I have become increasingly skeptical of the long-range enterprise as I graduated from starry-eyed student to professional historical linguist. It is a fact that the promise of demonstrating the genetic unity of all human languages originally attracted me to linguistics, but today the reality is that I now daily deal with the problems of determining whether two words are related when they belong to languages that are almost mutually intelligible.

Now having some years of experience in phonological and lexical reconstruction (in my case with the Bahnaric languages, a branch of Mon-Khmer), I have discovered for myself just how easy it is to find words in one language that resemble in form and meaning the words of another language, regardless of genetic relationship.

When it comes to obvious language families there are many cognates, and among these one can find distinctive sub-systems which identify the family, and usefully distinguish real cognates from phoney ones much of the time. Yet, even in well known language families, there must be some (if not many) apparently good etymologies which in fact are the result of chance alone, but are undetectable by any means because the chance resemblances mimic the patterns of the sub-systems which are inherited from proto-languages. However, we do not believe that families like Slavonic, Turkic or Katuic are artifacts of chance because it is obvious that chance would not have generated such large sets of resemblances in only one, two or three thousand years. But, when a supposed language family is around eight or ten thousand years, it is clear to me and many others that the resemblances that arise by chance alone (and may be highlighted by selective citation) can easily swamp those which are due to genetic inheritence—and an enthusiastic searcher can find among these bogus etymologies apparently non-trival sub-systems which are more complex and regular than those which are actually remnants of the proto-language.

I could offer any number of references to other scholars' views and their calculations, but I am sure that the readership of this journal (MT) are already well aware of those views and also have their own which they may or may not be ready to change. I make these remarks as much to clear the air and make my own views known as I do to prepare the reader for the text that follows.

So what are the chances for constructing a hypothesis of language relationship, complete with 'cognates' and grammatical parallels? There are various calculations around; the most recent example I have seen is Ringe (1999) on the chances of matching CVC roots (which is, much to my relief, light-years ahead of his earlier papers on the same theme). What follows is my own spin on the same topic, and I am sure that it is one that will have some of my dearest colleagues wondering just what has happened to their old friend. It is rather sloppier than Ringe's, but I have tried to model the real conduct of investigators rather than rigorous statistical models.

First of all it is apparent to me that, when one begins to investigate a possible remote relationship, the first step is to compile a list of similar morphemes, basically trying to match segments with the same (or similar) places and manners of articulation. This is exactly what I did in preparing my paper on the external relations of Ainu, which appeared in the last MT (Sidwell 1998). That paper actually appeared after a long delay (uncharacteristic for MT). My views have moved on somewhat since then. In that paper I attempted to use the notion of 'basic vocabulary' as a control on otherwise unconstrained resemblance-based comparisons, arguing that, if the rate of resemblance increases, as the comparisons are restricted to more basic vocabulary, a genetic explanation is more likely. It was a nice idea, but it was still an unbalanced approach, because there was no real test on the statistical significance of claimed resemblances. Given that so much of the long-range enterprise is based on finding formal agreement only at the level of the same place and manner of articulation (ignoring the fact that often even less agreement is offered), what are the chances of finding such agreements between two 100-word lists?

Let us conduct a thought experiment dealing with imaginary languages having morphemes with CVC structure. (Perhaps there are also simpler and more complex morphemes, but statistically the effect can be assumed to balance out). There will be stops, nasals, resonants and fricatives, and they will have five places of articulation. The set of final consonants is the same as the set of initials. We can effectively ignore the vowels—for the purpose of detecting distant relations they can count for nothing. (For the Bahnaric family I work with, there are hideously complicated vowel correspondences between languages which can only be a few hundred years apart).

Now, if we begin comparing morphemes between two languages, looking for segments that agree in manner and place of articulation, it is already clear that there will be some chance resemblances, because the total set of possible forms is limited. The total set consonants is be calculated as:

5 (places of articulation) x 4 (manners of articulation) = 20

Ignoring the vowels, we can calculate the number of potential acceptable CVC comparisons at:

20 x 20 = 400

But this is not correct, because there are always collocational restrictions, such as that initials and finals may not be permitted to have the same place and manner of articulation unless they are imitative or 'baby-talk'. Also, there will always be various combinations that simply do not occur, or occur very infrequently, so that effectively they can be removed from the equation. These restrictions may greatly reduce the number of available combinations. (English is rather odd in allowing virtually the same set of consonants both initially and finally.) Let us say that we have 200 possible combinations of initial and final consonants or hypothetical morphemes. I'll call them p-morphs (not an unrealistic simulation: e.g., Japanese has only 110 distinctive syllables (Taylor & Taylor 1995)).

If we take 2 random lists of 100 p-morphs, each drawn from the pool of 200, we expect to see about 50 matches between them, counting agreement in form alone. However, each p-morph has a semantic value, and these have to match as well to count as potential cognates. The semantic matches do not have to be perfect, because we are looking for genetic relationships (e.g., the words for 'hot' and 'cold' may have swapped meanings, but they are still cognate), but for the moment let us insist on exact translations. The chance of the p-morph on list *a* with a given meaning having the same form as the randomly generated p-morph of the same meaning on list *b* is:

0.5 x 100 = 0.005 or 1/200.

In other words, we expect that about half the time there will be one perfect match in form and meaning between our random lists. It doesn't sound like much, does it? It doesn't look anything like the supposed 29-41% cognacy claimed by Bengtson for Basque and Caucasian, and for good reason, because the thought experiment just conducted doesn't actually model what people do. So what really happens?

In the real world, investigators have before them wordlists of various lengths, often complete dictionaries of perhaps 50,000 words (for example, I do if I want to compare Chinese and Tibetan). For the moment, let us assume we have some lists of about 1000 words each, such as the popular vocabularies of unwritten languages published by the Summer Institute of Linguistics. Next, let us assume some semantic latitude, such that, for example, the term for any large tree may be compared to the name for any other large tree, or any small furry animal word can be compared to another small furry animal word, and so on. This effectively reduces the number of semantic distinctions on the 1000 word lists to, say, 200.

Next, let's look at the actual types of formal comparisons which are offered in texts such as John's letter which stimulated this discussion. In that letter one finds comparisons such as: bilabials with uvulars, alveolars with palatals, stops with resonants, nasals with laterals, various consonants with zero, to name a few. So even the restriction that segments must share manner and place is not respected. Actually it seems that an adjacent manner or place is enough, so that for example, John compares Basque *bi* ~ *bi-ga* 'two' with PNC **q'Hwǟ* 'two', presumably because /b/ and /w/ are both labial, even though they do not have the same manner of articulation.

Let us maintain the requirement that segments must agree at least in place of articulation. The effective number of consonants is no longer 20 but only five! This means that in practice there are only 25 distinct CVC p-morphs, rather than 400.

Now we begin the thought experiment again. We have two lists of 1000 p-morphs. There are 25 distinctive forms distributed over 200 meanings (or groups of synonyms—an average of five in each group). We take the p-morph at the top of list *a*, take note of the form and meaning, and begin to scan down list *b*, looking for a match. Naturally we look for semantically similar items, and we find a group of five synonyms in the same semantic field. Any p-morph can have one of 25 distinct forms, so the chance is 5/25 or 1/5 that there will be an agreement in form, comparing one to a list of five. However, it does not mean that we will find a total of 5/1000 or 200 matches. In practice we will be happy to find the best match from each group of five synonyms on list *a* with one of five synonyms on list *b*. This will happen about 1/5 of the time, so in total we can expect to find at about 40 matches (1/5 x 1/5 x 1000 = 40), many of which will be very impressive. This is already enough to swamp any real signal that might remain from a genuine but very distant genetic relationship between two languages.

Using more and longer lists, even if formal and semantic criteria are tightened somewhat, one can proceed to compile hundreds of comparisons, establishing regular patterns that mimic the kinds of patterning attributable to genetic processes. By chance alone many of these resemblances will be spectacularly good, and this will give some investigators the kind of positive reinforcement that quickly builds their convictions, ultimately forging a strong emotional attachment to the idea that a real thing has been discovered. Now the problem is not just statistical, it is psychological. Thus it will take far more than hard evidence and logic to untangle.

I took a look back at that Summer 1994-95 (that's Southern Hemisphere summer) issue of *Dhumbadji!* (now *History of Language*), with the long paper by Trask that stimulated the hue and cry which filled the bulk of Mother Tongue 1 (1995). On pages five and six we are treated to Trask's tongue-in-cheek comparison of Basque and Hungarian:

> I have chosen Hungarian [....] because I happen to have a large Hungarian-English dictionary at hand, and because I do not suffer from the inconvenient handicap of knowing anything about the history of Hungarian. My list of 65 Basque-Hungarian resemblances is rather longer than some I have seen presented in defense of other proposed links, but it took me somewhat less than four hours to assemble.

That list looks absurd, even to long-rangers, because *we all know* that Basque and Hungarian share no special relationship. But when we see the same done for Basque and another language whose genetic relations are not clear, we don't know what to make of it, because we don't already have a counter hypothesis. However, in both cases the level of evidence is similar, so a dispassionate eye must find them both equally absurd.

So do I conclude that long range linguistics is not possible? Not quite, and in fact I hope that it doesn't come to that. After all, it was the burst of interest in long range linguistics of the late 1980s that got me into linguistics (and not, say, archaeology, my first love), and still underlies my continuing deep respect for many members of the long range community. It's just that I'm not quite as starry-eyed as I used to be.

Dr Paul Sidwell
Department of Linguistics (Faculty of Arts)
Australian National University
ACT, 0200, Australia

References:

Ringe, Donald. 1999. How hard is it to match CVC-roots? *Transactions of the Philological Society* 97.2:213-244.

Sidwell, Paul. 1998. The External Relations of Ainu: Problems and Prospects. *Mother Tongue* 4:33-39.

Taylor, Insup & Martin Taylor. 1995. *Writing and Literacy in Chinese, Korean and Japanese*. Amsterdam/Philadelphia, John Benjamins Publishing Company.

Trask, Larry. 1994-5. Basque the search for relatives. *Dhumbadji!* 2.1:3-54.

Trask, Larry. 1995. Basque and Dene-Caucasian: A critique from the Basque side. *Mother Tongue* 1:3-82.

Comments on Bengtson's Basque-Caucasian Comparisons

By R.L. Trask
University of Sussex

[Editor's note: Trask had access to an earlier version of Bengtson's paper, so that comparison numbers and chapter headings referred to by Trask may be different from those in Bengtson's final version.]

1. 'bird'. Basque _txori ~ xori_ cannot possibly be the earliest form of the word, since palatal consonants do not occur in native and ancient lexical items, but only in expressive variants of these. As I may have pointed out before, the earlier word for 'bird' was *_zori_, a word whose primary form developed the transferred sense of 'omen' and then 'luck, fortune', while its expressive variant _(t)xori_ retained the sense of 'bird'. It is therefore *_zori_ which must be the Basque comparandum.

2. 'bone'. Though _(h)ezur_ is today the most widespread form, the Bizkaian variant _azur_ and the Roncalese variants _ẽzur_ and _enzur_ require us to reconstruct *_enazur_.

3. ' breast'. Basque **_bular_** is the usual word for 'chest' in most varieties, and this appears to be its original meaning. True, it is also widely used for 'female breast', and it is specialized in this sense in Gipuzkoan, but the evidence suggests that this – together with 'mother's milk' – is an extended sense. After all, Gipuzkoan is a central dialect, and we know what you think of forms found in central dialects. For 'chest', Gipuzkoan uses _kolko_ and _papar_, both words whose forms show that they are not ancient.

Moreover, alongside **_bular_**, we also find the variants **_burar_** and **_budar_**. This kind of variation is very familiar, and it points clearly to an original *_**burar**_, with the usual Basque dissimilation of /r/ to any of /l/, /d/ or zero in the configuration /rVrr/. This conclusion is reinforced by the observation that Zuberoan also has **_bular_** (or **_bolar_**). An original *_**bular**_ should have yielded *_**bülar**_ in Zuberoan, since original /u/ was fronted in that dialect in most circumstances, though not before the tap /r/. Hence the Zuberoan form likewise requires *_**burar**_.

4. 'cold'. Basque _(h)otz_ 'cold' is universal and seemingly ancient.

5. 'come'. Though it is entirely confined to the northeastern part of the country, _jaugin_ 'come', with its contracted form _jin_, appears to be ancient in the language.

6. 'day'. Basque **_egun_** 'day' is universal and seemingly ancient.

7. 'die'. Basque **_hil_** is everywhere the ordinary adjective for 'dead'. It is also the ordinary verb for 'die', and, in most varieties, the ordinary word for 'kill'. But the word has a uniquely anomalous form for a verb: in fact, it looks nothing like a verb at all, while its form as an adjective is unremarkable. Accordingly, we may conclude that the adjectival use is the oldest, and that the verbal uses are secondary. Any earlier verb for 'die' has seemingly been lost

without trace, but an earlier verb for 'kill', *ero*, is well attested in early texts and still alive today in the east. Accordingly, this word is only available for comparison as an adjective.

8. 'dirty'. Basque *zikin* 'dirty' is close to universal in the language and seemingly ancient.

9. 'dog'. The word *(h)or*, with its curious combining form *ora-*, is strictly eastern, except that a form *ora*, of uncertain reality, is barely recorded in Old Bizkaian. The variant with /h/ is strictly Zuberoan, and Zuberoan is the dialect which routinely adds the aspiration to all monosyllables which can bear it, so **or* is probably the etymon. There seems no reason to doubt the antiquity of the word, but it is perhaps too short and formless to serve usefully as a comparandum.

10. 'dry'. Among the variants of this word, *eihar* ~ *ihar* ~ *igar*, it seems clear that *eihar* is the most conservative, since the other forms can be readily derived from it by familiar processes.

11. 'ear'. The Basque word has the variants *belarri* ~ *beharri* ~ *begarri* ~ *biarri* (at least). It is impossible for *belarri* to be the most conservative form, since this cannot explain the other variants. Almost certainly the form to be reconstructed is **berarri*, again with the usual dissimilation of /r/ in the configuration /rVrr/.

12. 'earth'. Basque *lur* is universal and seemingly ancient.

13. 'eye'. Basque *begi* is universal and seemingly ancient.

14. 'fall' [1]. Basque *erori* has the apparent root *-ror-*, which is the only thing that should be compared, since the prefix *e-* and the suffix *-i* recur constantly in ancient verbs. This verb has a root of unusual shape: almost no other Basque root or stem contains two instances of the tapped /r/.

15. 'fall' [2]. The verb *jausi* is strictly Bizkaian in the sense of 'fall', all other varieties having only *erori*. The root is apparently *-aus-*. There may be a problem with the semantics. One early Bizkaian text uses the verb in the sense of 'surprise', and an identical verb is attested in the east in the sense of 'consent, agree, acquiesce'.

16. 'far'. The word *urrun* is found in both the east and the west, though often in the variant form *urrin*. The center of the country has instead the apparent derivative *urruti*, which may contain the familiar adjective-forming suffix *-ti*. A connection has been suggested with *(h)urbil* 'near', but the relation, if any,

is obscure. However, see the comments under item 96 below.

17. 'fat/grease'. The presence of nasal vowels in the eastern derivatives of *ziho* points unmistakably to an original **zino.*

18. 'few'. Today *guti*, with its expressive variants *gutti* and *gutxi*, is a quantifier meaning 'few, little, not many, not much'. But, in the medieval period (it is attested then) and in the 16th century, it was an adjective meaning 'small'. The word has an anomalous form for a monomorphemic lexical item. I've discussed it with several colleagues, and we all suspect that the word contains the common adjective-forming suffix *-ti* attached to an unknown stem.

19. 'fire'. Basque *su* is the universal word for 'fire', but the word is rather short and formless to be used as a comparandum.

20. 'fly'. For some reason, you've chosen the verb *hegaz egin*. But this is merely a typical Basque compound verb formed with the light verb *egin* 'do, make'. The first item here is *hegaz* ' by wing', consisting of *hega* 'wing' plus the instrumental suffix *-z*. The whole thing is literally 'make with the wings'. The only item available for comparison here is the noun *(h)ega* 'wing', which has curious variants *(h)ego* and *(h)egal*. The odd variation illustrated is unique to this word, and puzzling.

21. 'fog'. The Basque 'fog' words are a famous and messy problem. First, *lanbro*, which commonly means 'dense fog', 'pea-souper', cannot possibly be ancient in the language, with that /br/ cluster: plosive-liquid clusters were prohibited in Pre-Basque and were always eliminated in borrowings from Latin and early Romance. The word may be either a contraction or a borrowing, and both have been suggested. Second, it is scarcely likely that these two represent the same word, or even the same stem: they appear to be unrelated. Third, both words exhibit a range of other senses, such as 'mist, spray, spume, steam, vapour, cloud, drizzle' and also 'a certain disease of wheat', 'nearsighted' and 'cataract (of the eye)'. This makes determination of the original senses difficult and doubtful. Fourth, at least for *lanbro*, there is a possible Gascon source. All this is laid out in Agud and Tovar's etymological dictionary; the entries are too long for me to reproduce here.

22. 'foot'. Basque *oin* is universal and seemingly ancient. Given the absence of a native word for 'leg', this word may once have meant 'leg' as well as 'foot'.

23. 'four'. There is no doubt that *laur* is the most conservative form of the word: in all varieties, the stem is *laur-* when a vowel-initial suffix is added.

24. 'freeze'. There is no point in citing the verb *izoztu*, which is a transparent derivative of the noun *izotz* 'frost, ice'. This in turn is suspected by everyone of being a compound whose second member is *(h)otz* 'cold', even though the first element is unrecognizable. It might be *ihintz* ~ *intz* 'dew'.

25. 'guts'. It is certain that eastern *(h)ertze* is more conservative than western *este*, which results from a familiar phonological development.

26. 'hair' [1]. The Basque word *ile* exhibits a surprisingly large number of regional variants, making it difficult to reconstruct the Pre-Basque form. But we can be certain that it did not contain a simple /l/, since this would have developed regularly into /r/, which is found nowhere.

27. 'hair' [2]. Basque *bilo* cannot derive from an ancestral form with /l/, since, once again, this would have developed into /r/, found nowhere. Michelena once made the ingenious suggestion that this word might derive from a cross between Latin *pilum* 'a single hair' and Latin *villum* 'tuft of hair', with the meaning of the first but the form of the second. Don't know if this is right or not.

28. 'hand'. Basque *esku* is universal and seemingly ancient.

29. 'head'. Basque *buru* is universal and seemingly ancient. But *mutur* 'snout' is quite otherwise. This has an impossible form for an ancient lexical item, and it is unquestionably an expressive formation of a familiar kind. Formations of the type *mV(S)PVR*, where *P* is /t/, /k/ or /tz/, *R* is /r/ or /l/, and the optional *S* is any sibilant, are very frequent as expressive coinages. Moreover, words denoting projections are commonly created expressively in Basque. Unlike the vast majority of expressive formations, this one is found throughout the language, or nearly so, except that in some areas it appears in the variant form *mustur* — another indication of expressive origin, since ordinary lexical items do not exhibit this kind of variation.

30. 'hear'. Basque *entzun*, with its anomalous form, appears to be a contraction of something longer, possibly **enezun*.

31. 'hit'. Basque *jo* is universal and probably ancient, but an apparent root *-o-* is about as formless as you can get.

32. 'horn'. Basque *adar* is universal and seemingly ancient, but the existence of the very similar Celtic word represented by Old Irish *adarc* 'horn' has induced many Vasconists to see this as a loan from Celtic.

33. 'how'. There is no point in treating *nola* and *zelan* as independent

comparanda, since they transparently consist of the interrogative stems *no-* and *ze-* plus the manner-adverbial suffix *-la(n)*, found also in other formations.

34. 'I'. Basque *ni* is universal and undoubtedly ancient, though there are grounds for suspecting an ancient stem alternation **ni/*da*.

35. 'knee'. Basque *belaun* is universal and seemingly ancient, but it cannot derive from an ancestral form with /l/, since this would have changed to /r/. The combining form *belaur-* is regular for a native noun ending in /n/, and cannot be adduced in comparisons.

36. 'know (a fact)'. To be precise, Basque *jakin* means 'know (a fact)' in the imperfective, but 'find out' in the perfective. It is possible that the word derives from *jan* 'eat', much as the Latin verb *sapere* 'taste' developed into 'know' in western Romance.

37. 'know (a person)'. To be precise, *ezagutu* means 'recognize, be acquainted with, know (a person)' in the imperfective, but 'become familiar with, meet (a person)' in the perfective. The presence of the suffix *-tu*, borrowed from Latin, shows that the verb cannot be ancient in this form. However, it is highly likely that this was originally of the form **ezagun*. The word *ezagun* still exists today, but only as an adjective meaning 'familiar, well-known'. In all likelihood, the old participle of this verb became specialized as an adjective after the creation of the new participle in *-tu*, as appears to have happened in a number of other cases, and exactly as has happened in English: ' I have mowed the lawn' but ' new-mown hay', ' the metal has melted' but 'molten metal', and so on.

38. 'leg' [1]. It is scarcely likely that *zango* is native, in spite of its persuasively native form. Michelena derived it firmly from the Romance word which appears in Castilian as *zanco* 'stilt' (and other senses). Anyway, both the form *zango* and the sense 'leg' are confined to part of the east. Elsewhere the form is *zanko* (an impossible form for a native word), and the sense is variously 'calf (of the leg)', 'paw (of an animal)', or 'stalk (of a plant)'.

39. 'leg' [2]. And *hanka* is certainly not native: plosives were categorically voiced after /n/ in the medieval period in all but the eastern dialects, and so the /nk/ cluster here reveals a late origin. And that origin is obvious: it is the Germanic (probably Frankish) word **hanka* 'haunch', borrowed widely into western Romance. The Germanic word still exists today, as *hanke*, in Dutch and Low German, and the Romance borrowing is continued in Spanish and Italian *anca* 'haunch' and (Old and modern) French *hanche*, the source of English *haunch*. The Romance word was borrowed into Basque, where it means 'haunch' in the French Basque Country, but variously 'rump, buttocks',

‘leg, foot’ or ‘paw’ south of the Pyrenees.

40. ‘male’. Basque _ar_ is probably ancient, but it is too short and formless to serve as a comparandum. About half the languages on the planet seem to have a similar-looking word for ‘male’ or ‘man’: see the impressive list collected in Agud and Tovar’s dictionary.

41. ‘many’. The word _asko_ means ‘many’ only in the west. In the east, it means ‘enough’, just like the probably related eastern _aski_. Moreover, in the east, the word is preposed to its head, just like a _-ko_ phrase, suggesting that it was originally a derivative in _-ko_. I think everybody agrees that the eastern sense is conservative, that the word is a derivative in _-ko_, and that the likely source for both _asko_ and _aski_ is the curious verb _ase_ ‘be satiated, be satisfied’.

42. ‘moon’. You appear to be explicitly taking Basque _argi_ for comparison here, but this is wrong: _argi_ is merely the ordinary Basque word for ‘light’, and it has nothing to do with ‘moon’. The original Basque word for ‘moon’ is not recorded, but we can reconstruct it as _*iLV_, probably _*iLe_, where _L_ represents Michelena’s fortis lateral, probably a long or geminated lateral */ll/. It appears that this word originally meant both ‘ month’ and ‘moon’, but that then the Basques chose to distinguish these two by forming compounds. For ‘moon’, the most widespread form is _(h)ilargi_, literally ‘moonlight’. For ‘month’, the universal word is _(h)ilabete_, literally ‘full moon’, from _bete_ ‘full’ (western _ilbete_ apparently still means only ‘full moon’, and not ‘month’). One northern variety has the different, and puzzling, formation _ilaski_. (Azkue reports this, wrongly, as _*ilazki_, and tries to relate it to _eguzki_ ‘sun’.) The eastern word _argizagi_ is entirely different in formation: it combines _argi_ ‘light’ with an element _-zagi_, possibly the same element which occurs in _buruzagi_ ‘chief’ (from _buru_ ‘head’); this may be the same morpheme as _*-zani_ ‘guardian’, modern _zai ˜ zain_.

This original _*iLV_ can still be found inside various formations, such as western _ilgora_ ‘fourth quarter of the moon’. It also occurs in several names of months, such as _ilbeltz_ ‘January’ (_beltz_ ‘black’), _garagarril_ ‘June, July’ (_garagar_ ‘barley’), _uztail_ ‘July’ (_uzta_ ‘grain’, ‘harvest’), and _urril_ ‘October’ (_urre_ ‘gold’ ?).

Consequently, only _*iLV_ is available for comparison in the sense of ‘moon’.

43. ‘narrow’ [1]. Well, _estu_ is only the western word. The eastern word is _hertsi_, and we think these are the same word in origin. Since _hertsi_ is also a verb meaning ‘squeeze, bind, restrict, close’, we think that the participle of this verb came into use as an adjective meaning ‘tight, narrow, severe’, while western varieties transferred the verb from the native _-i_ class to the new _-tu_ class, yielding a participle (and adjective) _*erstu_ > _estu_. All this makes

perfect sense, but note that the form of the original verb *hertsi* requires it to be derived from a nominal or adjectival stem **herts-*, of unknown meaning. And this stem alone, whatever its original sense, is all that is available for comparison: all the other forms are readily derivable within Basque.

44. 'narrow' [2]. Various problems here. The word *mehar* does indeed mean 'narrow', but it is a transparent extension of the widespread word *mehe* 'thin, slender'. And this has nasal vowels in the east, obliging us to reconstruct **bene*, with the usual developments. North of the Pyrenees, we find only *mehar*; south, we find both *mear* and *medar*, with the /d/ clearly inserted to break up the hiatus, a common process in Basque.

45. 'night'. The universal Basque word is *gau*, which appears to be ancient. The variant stem *gab-* is secondary and confined to western varieties, where it arises when *gau* is followed by a vowel, as in *gaua* 'the night' (western *gaba*) and *gau on* 'good night' (western *gabon*). The variant stem *gab-* is not available for comparison.

46. 'not'. Basque *ez* is universal, except that Bizkaian exhibits a variant *ze* in specified circumstances. We are probably looking at the remnants of some ancient inflection, perhaps a negative verb.

47. 'old'. Basque *zahar* is universal and assuredly ancient, since it is rather clearly attested in Aquitanian as *SA.HAR*.

48. 'root'. Basque *erro* is probably today the most widespread of the several words for 'root'. However, two early writers testify that the word once meant, not 'root', but 'rhizome' — that is, a horizontal underground stem which sends up new shoots at intervals.

49. 'round'. Basque *biribil* ~ *borobil*, with its impossible shape, is clearly a reduplication of the ancient item **bil* 'round', attested nowhere as an independent word but of frequent occurrence in compounds, as in *ukabil* 'fist' (*uko* 'forearm') and in *gurpil* 'cartwheel' (*gurdi* 'cart'). Only this **bil* is available for comparison.

50. 'sand'. Basque *(h)ondar* means both 'sand' and 'remains, residue'; these are probably the same word, but we can't be certain. It is possible that the word contains the hypothetical collective suffix *-*ar*, if this is real.

51. 'see'. Basque *ikusi* is universal and surely ancient, but it derives from *ekusi* (attested), and only the root *-kus-* is available for comparison.

52. 'sew'. Basque *josi* is universal and seemingly ancient; its root is

apparently *-os-*, again rather short and formless.

53. 'sister'. The removal of the familiar kinship-term suffix *-ba* is justified, but removal of the initial /a/ is not. Anyway, since the eastern dialects have nasal vowels in this word, we may safely reconstruct **aniz-ba*. This is one of several kinship terms with initial *an-* or **aN-*; we don't know if there is a connection or not.

54. 'sit'. Well, *jarri* doesn't really mean 'sit'. Its general sense is 'put, place'. With a complement, it often means 'put oneself (into some position)'. In the west, it has acquired a secondary sense of 'sit down'. But 'sit' is not lexicalized in Basque.

55. 'skin'. Apart from the Bizkaian variant *narru*, Basque *larru* is universal and presumably ancient.

56. 'sky'. No, no chance. Basque *zeru* unquestionably derives from some Romance reflex of Latin *caelum*, of the approximate form **tselu*. This is completely unproblematical, and it is supported by the existence of the Zuberoan variant *zelü* (Zuberoan, alone among the dialects, does not always undergo the otherwise categorical shift of intervocalic /l/ to /ɾ/). The curious Roncalese variant *zeuru* ~ *zeuri* is not a problem: this odd-looking phonological development is well-known in Roncalese. For example, common *diru* 'money' (Zuberoan *diharü*), which everybody agrees is from Latin *denarium*, appears in Roncalese as *deuri*, with the same development.

The native Basque word for 'sky' is probably **ortzi*. This word is nowhere attested as an independent word in this sense, though it is attested in related senses like 'daylight, bright sky'; however, it occurs very widely in compounds, such as *ortzadar* 'rainbow' (*adar* 'horn'). The derivatives of **ortzi* are numerous, widespread and attested very early; in contrast, derivatives of *zeru* are very few and almost never attested before the 19th or 20th century. QED, I'd say.

57. 'sleep'. Once again, there is no point in citing the compound verb. The noun here is the universal *lo* 'sleep', surely ancient but again rather short and formless.

58. 'small'. Basque *tipi* ~ *tiki*, with its extraordinary form (it violates at least three morpheme-structure constraints applying to native words) is, if not a loan, assuredly an expressive formation, on a par with English 'teensy'. This expressive formation is not available for comparison: such formations are far too common in languages. Look at the numerous Romance formations like Castilian *chico* 'small', French *petit* 'small' and Italian *piccolo* 'small', which have no Latin sources.

59. 'snow'. The Basque word is variously *elur* ~ *erur* ~ *edur*. As usual in these cases, the /l/ cannot be original, and we must reconstruct **erur*, with the usual dissimilation of /r/ in the configuration /rVrr/.

60. 'spit'. Again, you've cited the compound verb instead of the bare noun. Now, the most widespread Basque word for 'spit' is *tu*, with expressive variants *ttu* and *txu* and a Gipuzkoan variant *to*. This is indisputably of imitative origin: words for 'spit' of the form /tu-/ are widespread in the world's languages, and anyway you only have to listen to a French Basque pronounce the word, with his aspirated /t/, to realize the imitative nature of the word.

The form *listu* is confined to High Navarrese, and it appears to be a compound of *tu*. I think there's little doubt that the first element is the Navarrese word *lits* ' small spot of liquid on the ground'. Presumably this word originally meant ' gob of spit on the ground', and was then generalized to 'spit'.

Note also that several other varieties of Basque have a word *txistu* for 'spit'; this appears to be an expressive variant of *listu*. And this *txistu* is also the name of the vernacular end-blown flute typical of the Basque Country. If you've ever played a flute, you'll recognize the connection between flutes and saliva.

61. 'squeeze'. As I explained earlier, the verb *hertsi* is the same word that you cited as *estu* in number 42. And it's really rather naughty of you to cite two regional variants of a single word in two entirely different comparison sets. To my mind, this only shows the futility of trying to work with superficial resemblances.

62. 'star'. Basque *izar* is universal and doubtless ancient, but it too may perhaps contain the hypothetical collective suffix **-ar*, if this is real.

63. 'stick'. I think everybody agrees that Basque *makila* derives from Latin *baccilla* ' rods, wands'. With its initial /m/, the word looks anything but native.

64. 'stone'. Basque *(h)arri* is universal and doubtless ancient, though there have been strenuous attempts at deriving it from a putative stem **karr-* found in Romance (and possibly, but dubiously, in Celtic) and attributed to some unknown "pre-Roman" language.

65. 'sun'. The Basque word is a transparent derivative of *egun* 'day', and you are correct not to treat it as a separate comparandum.

66. 'tail'. Basque *buztan* is universal and seemingly ancient.

67. 'that'. Mesial *hori* and distal *hura* should not be conflated. The first is pretty clearly ancient. But the second is pretty clearly not. The original distal demonstrative was **(h)ar*, which survives today as the oblique singular stem *(h)ar-* of *(h)ura*, as the Bizkaian distal *a* (Bizkaian does not have the innovating *(h)ura*), and as the universal definite article *-a*, oblique singular stem *-ar-*. The origin of the new form *(h)ura* is obscure, though the late Alfonso Irigoyen some years ago published a very interesting proposal which is too complex for me to reproduce here.

68. 'there'. Again, mesial *(h)or* and distal *(h)an* should not be conflated. The first form is isolated and mysterious, since it alone, among all the locative forms in the language, fails to exhibit the locative ending *-n*. But *(h)an* is nothing but the distal demonstrative stem plus locative *-n*.

69. 'this'. There is little doubt that eastern *haur* is more conservative than western *au*.

70. 'thou' [1]. It is beyond dispute that *(h)i* was originally the ordinary second-singular pronoun.

71. 'thou' [2]. But *zu*, as you note, was historically the second-plural pronoun, still in use as such in Bizkaian in the late 19th century.

72. 'tongue'. Well, how many times have we had this conversation? The Pre-Basque word was **bini* and nothing else: this form alone can account for all of the numerous regional variants and combining forms, while **mihi* has no chance at all.

73. 'tree'. There are about ten different Basque words attested for 'tree', and this meaning seems to have been highly unstable in the language. The most widespread modern word, *zuhaitz*, is of uncertain but debatable antiquity (it is not recorded before 1630, and earlier texts use other words). In any case, as you note, this is pretty clearly a compound of *zur* 'wood', which alone should be the comparandum.

74. 'turn'. Basque *itzuli* is all but universal, being absent from the east, but doubtless of some antiquity. Its root is *-tzul-*, which alone is available for comparison. But the semantics may be a problem: the word has numerous senses, of which the earliest recorded is 'return, go back', rather than 'turn, rotate'.

75. 'twenty'. Basque *(h)ogei* is universal, apart from its eastern variant *hogoi*, and surely ancient.

76. ‘two’. It is clear that eastern *biga* is the more conservative form, and that *bi* originated as a contextual reduction of this.

77. ‘warm’. Basque *bero* is universal and doubtless ancient, but there’s a problem. Imitative forms of the type *ber(ber)* are very widespread as onomatopoeias for the sound of boiling, and these often acquire transferred senses like ‘ hot, burn, fire, cook’. Hence words of this type cannot be assigned full value in comparisons, since they are so often created independently.

78. ‘water’. Basque *ur* is universal and surely ancient, but again it is too short and formless to be of much value as a comparandum.

79. ‘what’. The stem is *ze-*, and *zer* can only be invoked separately if a match is proposed for the mysterious /r/, which does not appear to be the case here.

80. ‘when’. Again, the stem is *no-*, and there is no point in citing *noiz* separately unless a match is proposed for the puzzling ending, which again does not appear to be the case.

81. ‘where’. Same remark again, except that this time *non* plainly ends in the ordinary locative *-n*, which requires no explanation.

82. ‘white’. Basque *zuri* is universal, except that some varieties use the originally expressive variant *(t)xuri*, and seemingly ancient. But there might be a problem. We know that Basque formerly possessed an adjective-forming suffix *-i*; this is no longer productive but still clearly visible in a few formations. But the language also has lots of adjectives ending in *-i* for which no obvious derivation is visible. Now, *zuri* is one of these, and Azkue once suggested a derivative of *zur* ‘wood’, which is phonologically perfect and semantically reasonable: the original sense would then have been something like ‘light-colored’. Don’t know if this is right, but it must be taken seriously.

83. ‘who’. The usual comment with interrogatives: the stem is familiar, but what is that Basque /r/? A comparison that could provide a compelling answer to this question would be more persuasive than one that can’t. This is an example of what I mean when I complain that the numerous comparisons of Basque with language L *never* shed any light on our problems.

84. ‘wing’. I’ve already discussed this under item 20 above. Once again, you are taking two variants of a single Basque word and comparing them separately with different Caucasian items which they happen to resemble separately. And this is out of order.

You really want to impress me? Than show me some good Caucasian data explaining clearly why the Basque word for 'wing' exhibits such curious variant forms. Now *that* would impress me.

85, 86. 'woman' [1], [2]. No. Consider some facts. From the earliest records until the 18th century, the usual Basque word for 'woman' was *emazte* (or its extended form *emazteki*). This is still the word for 'woman' in the Salazarese dialect today, but elsewhere *emazte* has been specialized to 'wife'. Its place has been taken by *emakume*, a word which is recorded at least once in the 17th century (and possibly only once), but which since the 18th century has been the ordinary word for 'woman' in all varieties except Salazarese, where it interestingly means 'girl'.

Now, comparison of *emazte* and *emakume* leads immediately to the observation that the shared element is the initial *ema-*, which we must therefore suspect of being the semantic center of both words — yet this is exactly the portion which you tear off and throw away without comment in both cases.

There's more. The universal Basque word for 'female' is *eme* — and this, of course, has the fully regular combining form *ema-*, as in *emamintza* 'hymen' (*mintza* 'membrane'), *emazurtz* 'female orphan' (*zurtz* 'orphan'), *ematxar* 'woman of ill repute' (*txar* 'bad') and *emabide* 'vagina' (*bide* 'way'). So, already it is very difficult to avoid the conclusion that both *emazte* and *emakume* are built upon *eme* 'female', and that the meaning derives primarily from this element, and not, as you assume, entirely from the second elements.

Now, that's perhaps already enough, but in fact we have good etymologies for both words. For *emazte*, whose second element is not transparent, Michelena proposed **ema-gazte* 'young female' (*gazte* 'young'). The loss of /g/ between low vowels is a sporadic but familiar process in Basque: for example, French *magasin* 'department store' occurs in French Basque as *maasin*, with loss of the /g/. Now, **emagazte*, being four syllables long and of high discourse frequency, might readily have undergone precisely this process, producing first **emaazte* and then the observed *emazte*. Not certain, but very plausible, and I think it's right, though it suggests a meaning shift from an unattested 'girl' to 'woman'. As for *emakume*, that's transparent: the second element is the very common suffix *-kume* 'offspring', found in numerous formationa like *katakume* 'kitten' (*katu* 'cat'). This is consistent with an original sense of 'girl', as attested in the seemingly conservative Salazarese dialect.

So, we have a virtually perfect account of both words: both are formed within Basque, by familiar patterns, with familiar elements and familiar phonological developments. Your account, in great contrast, obliges us to regard all this as a fantastic collection of astounding but meaningless coincidences: the ancient Basques, in your account, chose to stick the meaningless sequence *ema-* at the beginnings of both words for 'woman', and this meaningless sequence just happens, entirely by chance, to be identical with the combining form of the ordinary word for 'female'. Sorry, but I don't think so.

Finally, I might add that Basque *eme*, with its non-native form, is a transparent borrowing from Gascon *hemne* or a related Occitan form, the source being Latin *feminam* 'woman'.

87, 88. 'woods' [1], [2]. It is true that both *baso* and *oihan* usually mean 'woods, forest' today, though neither is found throughout the country. But there is some doubt about their earlier senses. Old Bizkaian *oian* means 'desert', and the combining form *basa-* of *baso* commonly means ' wild', ' wilderness', 'uncultivated land'. It is not clear that the presence of trees was always a requirement for the use of either of these words.

89. 'worm'. No. Bizkaian *aar* and Zuberoan *hãr* point clearly to original **anar*.

90. 'ye'. There is no point in citing this separately, since the modern plural *zuek* is transparently derived from earlier plural *zu*, as you point out elsewhere, by the addition of a plural suffix.

Addendum

91. 'bark'. Basque *azal* 'skin, bark', though not universal, is very widespread and presumably ancient, though the sense of 'bark' is perhaps more widespread than that of 'skin'.

92. 'belly'. Basque *sabel* conceivably contains the ancient morph **bel* 'dark', found in a sizeable number of formations. But this is far from certain, since the first element would be unidentifiable, and since it is not easy to think of a semantic value X for which 'dark X' would express 'stomach'.

93. 'claw'. Basque *atzamar* ~ *atzapar* ~ *aztapar* 'finger, claw', the last form being chiefly northern, is a familiar headache. It is difficult to tell which variant might be the most conservative. The one thing that is certain is that the word cannot be monomorphemic, since it has an impossible form for a monomorphemic word. The favourite guess is still a back-formation from *hatz (h)amarrak* 'the ten fingers', though not everyone is happy with this, and it admittedly has an anomalous word order. In any case, a root **-tapa-* is out of the question: Basque words do not have such forms.

94. 'go'. Basque *joan*, with an apparent but anomalous root *-oan-*, is universal and surely ancient. But there are grounds for suspecting a lost consonant between the /o/ and the /a/.

95. 'heavy'. Basque *astun* has been the subject of a good deal of discussion. Its form is consistent with ancient and monomorphemic status, but everyone who has looked at it has suspected the presence of the common adjective-forming suffix *-dun* 'having'. Moreover, there is a common word *astin* 'light, thin, insubstantial', and a synonymous Bizkaian word *astor*. These have led specialists to propose various connections among these words, ranging from mere contamination to a common stem. But nothing can be established, and so there can be no objection to citing *astun* as a comparandum.

96. 'near'. Basque *(h)urbil*, confined to part of the country, appears to be a derivative of a shorter word for 'near', found in *urrean* 'near', with the locative ending. The noun back-formed from this is *urre*, but we may be reasonably confident of an original **ur*, with familiar developments, also the source of *urren* 'nearest', with superlative *-en*. And this **ur* is again rather short and formless for comparison.

97. 'skin' [2]. See item 91.

98. 'tie'. The stem of the unusual verb *lotu* is *lot-*, and the absence of the prefix **e-* makes it clear that this stem is non-verbal in origin, assuming it is native, though an Occitan source has been suggested, not entirely persuasively.

Comments on "Recurrent regularities"

(1) No comment, except to note that the Basque trill /rr/ is *always* spelled <r> in any position other than between vowels.

(2) My only comment is that invoking "assimilation, dissimilation or metathesis" "in one language or the other" effectively allows you to match any two liquids at all.

(3) There is no evidence for **ardar* as the etymon of *adar* 'horn', and the possible Celtic match militates against this. As for **merdar*, this is indefensible: the etymon of Basque *mehar* ~ *medar* is **bene* + **-ar*.

(4) In my view, there can be no justification for regarding Basque /lh/ as a segment different from /l/. First, *all* occurrences of aspiration in Basque are subject to certain general constraints, constraints which cannot be adequately stated if some instances of the aspiration are regarded as distinct segments. Second, even in the aspirating dialects, there is considerable variation in the occurrence of /lh/: for example, we find *hilargi* ~ *ilhargi* 'moon', *hilerri* ~ *ilherri* ' cemetery', *alaba* ~ *alhaba* 'daughter', *zulo* ~ *zulho* ~ *zilho* 'hole', and many others. Third, even in obvious Romance loans, we find /lh/ alongside /l/, as in *mulo* ~ *mulho* 'stack', from late Latin **mulu(m)*, attested throughout

Gallo-Romance. Such forms would be inexplicable if the aspiration were anything other than a suprasegmental feature in origin. Fourth, the aspiration never prevents a preceding consonant from being intervocalic and developing in the normal way for intervocalic consonants – hardly possible if /h/ were a real consonant. Hence, for example, *bil(h)o* 'hair' cannot possibly derive from an original **bilho*, which would have developed to **bir(h)o*. And Basque *elur* ~ *elhur* 'snow', of course, cannot possibly have had an original lateral, since such a form would be incapable of accounting for the variant forms.

(5) Eastern *belhar* 'forehead' has, according to Agud and Tovar, an attested variant *belaar*, which points clearly to a bimorphemic origin.

(6) No new comments.

(7) Two of the three Basque comparanda here are unacceptable, for reasons outlined above.

Comments on "Morphological Parallels"

(a) On the alleged fossilized prefixes, I have nothing to add to my comments in earlier communications.

(b) I am staggered by the suggestion that Basque final *-r* is a "plural" marker whenever you can't match it. What on earth is a plural morpheme doing in words like *bizkar* 'back' and *belar* 'forehead'? In fact, it has long been noticed that Basque possesses a sizeable number of nouns all of which end in *-ar* and most of which denote things more commonly encountered in bunches than individually: *(h)ondar* 'sand', 'remains, residue'; *izar* 'star'; *ilar* ~ *idar* ' pea'; *negar* ~ *nigar* 'tears'; *sagar* 'apple'; and quite a few others. It has accordingly been suspected that this *-ar* (not *-r*) might represent a fossilized collective (not plural) suffix, but the evidence is not sufficient to draw such a conclusion.

By the way, Basque has not "generalized the plural ending in *-k*". The plural in *-k* occurs nowhere but in the three demonstratives and in the ordinary (or 'definite') article, which itself derives from a reduction of the distal demonstrative **har*. It has been extended to one or two other forms, like western *batzuk* 'some' (eastern *batzu*) and common *zuek* 'you (plural)', but that's about it. This marker is not found elsewhere in the language.

What would be required to prove a genetic link between Basque and other Eurasian language families

by Edward J. Vajda
Western Washington University
vajda@cc.wwu.edu

Readers of *Mother Tongue* need no extended introduction to the debate now swirling around proposals of genetic relationship between Basque and various other Eurasian linguistic islands and microfamilies (Trask 1995a). However, I think the basic concept of "proof" of genetic affinity between languages must itself be re-evaluated before commenting on John Bengtson's proposed Basque-Caucasian lexical parallels. This assessment comes from someone who is neither a Vascologist, nor a Caucasianist, nor even a "long-ranger" (a label I find infelicitous, since there is no precise division between comparativists investigating different time depths). Yet my lack of specialist credentials here might be defended as specifically useful, since in the field of linguistics it will be the large number of interested non-specialists like me who will ultimately decide the fate of any hypothesis of genetic relatedness between languages.

In practical terms, why is it that some proposed language families gain universal acceptance, while others, in the estimation of most linguists, never advance beyond the realm of mere speculation? What exactly has to happen before the validity of a given proposal is widely taken for granted by educated non-specialists? I don't believe that generations of linguists, like flocks of sheep, simply follow whoever happens to argue the loudest or sports the most impressive academic credentials. That sort of thing might happen occasionally, but eventually a hypothesis stands or falls depending on whether certain types of evidence have been placed before a broad audience in a clear and understandable way. There are many language families analogous in time depth to Indo-European for which this has been accomplished, and around which such consensus has developed. And there are more than a few generally accepted genetic groupings of considerably greater antiquity – Afroasiatic, to name one salient example. Many people today who take for granted the existence of these connections are not even aware who made the original proposal of genetic relatedness. Examining how these and other families have come to be generally accepted as valid genetic units might shed some light on what still needs to be accomplished by proponents of Dene-Caucasian or even Vasco-Caucasian, groupings at present accepted by only a few linguists. (The same can be said, by the way, regarding North Caucasian, a putative genetic unit comprising two generally accepted families: Abkhaz-Adyge and Nakh-Dagestanian.) Such an assessment is very relevant to *Mother Tongue* editor John Bengtson's request for peer review of his proposed lexical parallels involving Basque and various North Caucasian languages. My suggestions will in some ways mediate between the positions occupied by the so-called "lumpers" and "splitters" (two more epithets I regret, since they ignore the important contributions that each camp brings to genetic linguistic research). On the other hand, my ideas are intended to offer a distinct third view rather than a watered-down middle ground between the former two approaches.

First, I must concur with a position discussed repeatedly by Luigi Cavalli-Sforza, among others, that what is at issue is not whether two languages (or two populations) are or are not genetically related. The issue is the *degree* of putative genetic relatedness. I think it is safe to say that most linguists entertain the possibility that all human languages are most likely genetically related at some great time depth, but disagree on the issue of how or even whether this can be proven; so we can speak of putative degrees of relatedness in linguistics as well as biology. For example, it is not enough to assemble evidence indicating that Lithuanian and Irish are genetically related. One must also survey other languages to determine whether these two together form a valid genetic unit or belong to two separate subunits of a broader genetic grouping. This goal seems to be at the heart of Greenberg's "mass comparison" (or "multilateral comparison") approach, which probably should simply be called "genetic linguistic taxonomy." Because all human languages may in fact be related, the historical-comparative linguist must not only determine the fact of genetic relatedness itself, but must also demonstrate whether two or more related languages form a valid genetic unit *vis-à-vis* other known languages. Even if one accepts Bengtson's Basque/Caucasian parallels as evidence of descent from a common source, it must further be demonstrated that most of these parallels are unique to his proposed grouping (or to Dene-Caucasian) and do not represent "global etymologies" or any other kind of geographically widely dispersed lexical parallels. Most importantly, Bengtson must demonstrate that there is something about them that elevates them above the level of chance resemblances of the type which can be amply demonstrated between any two languages. This is very difficult to do without the support of some sort of systematicity in the data. Larry Trask's pseudo-inventory of Basque-English lexical parallels (Trask 1995b: 196-7; Trask 1996: 113) in basic vocabulary is instructive here. He obviously intended these lists to illustrate the difficulty inherent in determining whether Bengtson's Basque-Caucasian parallels are not equally spurious instances of coincidence. But at the same time, for the sake of argument, how can we be sure *a priori* that some of these or any other random lexical parallels are not shared by other languages besides the ones being compared, given that all languages may be ultimately genetically related? My point is this: a list of lexical parallels alone, regardless of the calculated percentage of basic vocabulary (based on the Swadesh list, the Yakhontov list or any other allegedly universal measure of lexical stability) is simply not enough to elevate a genetic hypothesis above the status of the logically competing coincidence hypothesis. Nor can it entirely rule out the "global etymology" hypothesis. If all languages descend from a common source, then perhaps some of the proposed Basque/Caucasian parallels are also to be found in branches of Nilotic, Afroasiatic, Dravidian, and so forth; consequently, their presence in the two groups being compared, even if taken as proof of genetic relationship, is not enough to elucidate the precise degree of relationship. Proof of genetic relatedness, then, boils down to assembling enough evidence to dispel every other competing hypothesis. And lexical data alone is simply not enough to do this in cases where parallel evidence from other aspects of linguistic structure are lacking.

This leads into my other point. Rather than define "proof" in linguistic investigations of genetic relatedness by proposing pseudo-specific percentages of various sorts, I would like to replace this notion with the fuzzier, but pragmatically more realistic, concept of "superior hypothesis." A proposal for a language family gains general acceptance if, and only if, two things happen: 1) enough parallels in language structure are found and described to permit rejection of alternative possible hypotheses such as coincidence or borrowing; and 2) it is demonstrated that other languages do not likewise share these or analogous features in ways that suggest a closer genetic association with any subset of the languages under investigation. (Obviously, most of the parallels between Irish and Lithuanian will be shared by Welsh, on the one hand, and Latvian, on the other—languages that must also be included in the taxonomic survey before any informed proposal can be made regarding the genetic relatedness of Irish to Lithuanian). Simply claiming (or even offering plausible evidence) that two languages are genetically related is, in and of itself, insufficient to clarify the genetic status of these languages. This factor becomes all the more important in cases where genetic affinity is proposed on the basis of random lexical parallels alone. Even if most of the parallels are presumed to be true cognates, what proof is there that something similar cannot be found in other languages as yet uninvestigated?

Two issues must be considered from the start by those creating or considering proposals of distant genetic relatedness: (1) the necessity of conducting a systematic taxonomy instead of isolated lexical comparisons between languages (and deep genetic comparisons involving Eurasian isolates such as Basque require investigating *all* Eurasian families equally and without preconceived conclusions, not just the few isolates and microfamilies which haven't been conclusively related to larger groupings); and (2) an understanding that we are always dealing with competing hypotheses rather than mathematical proofs.

So exactly what evidence, in addition to putative lexical parallels, is needed to build a hypothesis of genetic relatedness that supersedes all competing hypotheses? I would argue that the answer amounts to something less than a requirement for an all-encompassing set of systematic sound correspondences. Languages have been convincingly shown to be genetically related in the absence of such a complete set of proto-phonological data – Tlingit + Eyak-Athabaskan being a good example (though in this particular case, Leer 1990 has proposed a plausible sociohistorical source for the irregularity). Instead, the minimum that needs to be clearly demonstrated is the presence in all of the languages being compared of a *complex of shared structural features* which represent obvious genetic inheritance. In other words, their combination is diagnostic in the sense that it cannot plausibly be considered a mass of separate coincidences, and the features in question are too much a part of the grammar and core lexicon to have plausibly originated through language contact—though excluding contact is arguably more difficult in cases where languages are known to have been spoken contiguously for long periods. This complex of shared features will probably include at least some sound correspondences, demonstrable on the basis of cognates in basic vocabulary (with the word "basic" loosely defined), as well as core grammatical traits, such as shared characteristic patterns and constructs of morphosyntax (not isolated similarities between one or two inflectional affixes, which might easily be coincidental). I would argue that, in addition to shared vocabulary and their concomitant sound correspondences, each genetic linguistic grouping must share its own unique set of typological and structural traits. For Celtic these include a propensity for using stem initial consonant mutations, the presence of verb-initial word order, and inflected prepositions. While any one of these traits might be found by coincidence in other languages, their persistent *combination* is expressly unique to Celtic. For Tlingit-Eyak-Athabaskan, we have a unique type of verbal polysynthesis involving the interdigitation of derivational and inflectional categories superimposed upon a highly specific arrangement of prefixal positions. All branches of Afroasiatic likewise share diagnostic structural features, such as inflectional sound symbolism, as well as shared core vocabulary. Finally, it was in fact the characteristic verbal inflectional patterns shared by Latin, Greek and Sanskrit that gave rise to Sir William Jones's famous pronouncement, not an extensive set of "sound correspondences" (though the latter eventually came to light as well). The consensus for every widely accepted language family or genetic grouping rests on a similar core of evidence, even in the absence of more systematic sound correspondences that easily lend themselves to a detailed reconstruction of a proto-language. This is true of Yurok + Wiyot, as well as of Uralic + Yukaghir and many other groupings. Once the historical linguist has identified a diagnostic core of traits that cannot plausibly be coincidental, are too extensive to be explained by known language contact phenomena, and clearly do not exist together in any other languages, the hypothesis of genetic relatedness becomes the best explanation - not only for linguists living at the time the proposal is made, but also for future generations of scholars, each of whom will have to be quietly and individually convinced by the raw data rather than by the charisma or academic stature of the proposal's original supporters. Once this begins to happen, the hypothesis could be considered "proven," at least in practical terms. I can't help recalling the Sherlock Holmes approach to solving mysteries: "Once you have removed all other possible explanations, the one remaining, no matter how improbable, is the real answer." Many hypotheses of deep linguistic relatedness simply fail to remove the possibility of coincidence beyond possible doubt, at least for the majority of interested non-specialists who are exposed to the data. And it is this large group - the interested non-specialists - who will become the ultimate arbiters in the matter. The specialists may propose, but it is their future colleagues who will dispose. Ultimately, "Basque-Caucasian" will stand or fall based on the common sense intuition of many anonymous linguists who happen to be exposed to the evidence presented, and neither today's Basque or Caucasian specialists nor today's "long rangers" will ultimately have any additional influence on the fate of their pronouncements beyond the common-sense implications of the data they have succeeded in bringing to light.

Getting back to "Basque-Caucasian", I think Bengtson's lexical evidence alone, regardless of how it is presented or how forcefully it is argued, is simply not enough to elevate the genetic hypothesis above other logically competing explanations. Something qualitatively different is needed, as well. In addition to his lexical studies, Bengtson needs to demonstrate what is uniquely characteristic about the grammar and phonology of such a grouping as Basque-Caucasian. Is there a diagnostic combination of typological traits evinced by the languages in this grouping? If so, what is it? To begin finding this out, an extensive comparative study of the morphosyntax of all the languages in question must be conducted. This is no easy task, as this sort of comparative work requires considerable familiarity with the languages in question and not just a passing acquaintance with dictionary materials. In this connection, I believe Trask's list of structural features seemingly unique to Basque (1996, p.115-6) is an important and valuable gift to all those who hope to demonstrate external linguistic affinities for Basque. So far this list has remained without comment in the pages of *Mother Tongue*. In the final analysis, Basque will remain an isolate (or, as Bengtson has colorfully termed it, "an orphan forever") in the minds of most future linguists, until someone succeeds in providing plausible answers to at least a some of Trask's questions. At the present time I cannot be certain whether such a study will yield convincing evidence for a connection between Basque and any Caucasian language family (or even for a connection between Abkhaz-Adyge and Nakh-Dagestanian, whose most closely shared structural trait appears to be a general typological dissimilarity with respect to other Eurasian language groups rather than any genetically meaningful similarity). But regardless of the eventual consensus to come out of this sort of research, it must be done, and further genetic comparative work involving Basque and other language families should be encouraged. John Bengtson's pioneering efforts deserve praise both for extending the field of inquiry and for generating a forum of fresh debate on a number of fascinating issues.

References:

Leer, Jeff. 1990. "Tlingit: A portmanteau language family?" In *Linguistic Change and Reconstruction Methodology*, Philip Baldi, (ed.). (Trends in Linguistics. Studies and Monographs, 45.) Berlin and New York: Mouton de Gruyter. Pp. 73-98.

Trask, R.L. 1995a. "Basque and Dene-Caucasian: A critique from the Basque side." *Mother Tongue* I, Dec., 1995. Pp. 3-82.

Trask, R.L. 1995b. "Response to comments." *Mother Tongue* I, Dec., 1995. Pp. 172-201.

Trask, R.L. 1996. "Response to Starostin." *Mother Tongue* II, Dec., 1996. Pp. 111-17.

The Basque Language and its Closest Relatives

by Paul Whitehouse

Flat 3, Angel House, Pentonville Road,
London N1 9HJ
Tel: 0171-278 8180
Email: paul_whitehouse@talk21.com

I was flattered to be asked to contribute to the Basque Debate, particularly since I have no special knowledge of either Basque or its immediate relations. My only useful contribution would have been a global comparison along the lines of my recent Nihali/Kusunda article (*Mother Tongue III*), giving all the similarities I could find to Basque to see what light (if any) these shed on how Basque fits into the global picture. There was not time for this, however, either to make my database as wide as it needs to be or to collate the results into article form — and if there had been, the editor could not have allowed me space to publish it.

In lieu of this, then, I did a brief survey of that part of the database which is already collated, to get a rule-of-thumb, work-in-progress sense of where Basque fits into the grand scheme of things. I sought in the process to falsify both the Dene-Caucasic hypothesis and the Macro-Caucasic subgrouping within it by identifying better alternatives to them, in the form of either a single superior proposal or a spread of equally plausible proposals which would devalue the evidence for the Caucasic connection. I was unable to do either.

There was no other connection for which as many similarities could be found to Basque, either quantitively or qualititively. Conversely, the evidence for Basque-and-Caucasic was clearly superior to all other possible connections. There emphatically was *not* the equivalent body of evidence that should have been available everywhere I looked if the evidence linking Basque to Caucasic were nothing more than chance coincidence.

Let me be specific here. On the basis of the comparisons I have made it is clear that Basque has more in common with Dene-Caucasic than with each of the following:

Khoisan
Hadza
Sandawe
Kongo-Saharan [Gregerson]
Nilo-Saharan
Chari-Nile [Greenberg]
Songhay
Fur
Kunama
East Sudanic
Kadu
Kordofanian
Niger-Kordofanian
Niger-Congo
Mande [Westermann]
West Atlantic [Westermann]
Shabo
Ongota

Afrasian [Greenberg]
Nostratic [Illych-Svitich]
Sumerian
Amerind [Greenberg]
Proto-Eskimo-Aleutian [Mudrak]
Proto-Chukchi-Kamchatkan [Mudrak]
Nivkh
Japanese
Korean
Ainu
Macro-Australic [Bengtson]
Miao-Yao
Austo-Tai [Benedict]
Austric
Viet-Muong
Common Australian [Capell]
Paman
Tiwi
Nihali
Kusunda
Indo-Pacific
Great Andamanese
Little Andamanese
Tasmanian
Timor-Alor [Greenberg]
West New Guinea [Greenberg]
North New Guinea [Greenberg]
South New Guinea [Greenberg]
South West New Guinea [Greenberg]
East New Guinea [Greenberg]
Pacific [Greenberg]

I *think* this is everything I looked at first time round. I have still to investigate whether Basque might have some as-yet unrecognised special relationship to the various unlisted components of these larger phyla, but on the basis of what I *have* looked at this seems unlikely.

I am therefore obliged to accept, for want of a better alternative, that Basque is most closely related to Caucasic and Burushaski, and that this 'Macro-Caucasic' subgroup belongs in a larger Dene-Caucasic phylum with Yeniseian, Sino-Tibetan and Na-Dene. If anyone has found evidence for a more compelling relationship between Basque and any other genetic entity (particularly anything on the above list), please share it with us.

This proviso, 'for want of a better alternative', may sound like damning the proposals with faint praise. In fact, it is the essence of what we are trying to achieve. As Merritt Ruhlen keeps reminding us, the best that sciences like linguistics can aspire to, in the absence of absolute proof, is the best available interpretation of the data. Where there is a body of non-random similarities shared by a particular set of languages (in this case Dene-Caucasic, as outlined above), which cannot be matched by a body of non-random similarities between the languages in question and any other language(s), we are obliged to accept this as demonstrating a particular genetic relationship between them, *even if the degree of similarity is not ideal.*

This begs the question 'when may similarities be considered non-random?' In the specific case of Dene-Caucasic, I would answer that the similarities between Basque and the other Dene-Caucasic languages are non-random because they are significantly more numerous than those between Basque and non-Dene-Caucasic languages, and because John Bengtson is able to point to specific recurrent similarities which are definitely not evident elsewhere.

Bengtson goes further than this, calling the relationship 'obvious'. This is not a word I would ever use in the context of Proto-Caucasic reconstructions. Even after I had written them out using a more familiar phonetic transcription, I still had trouble with them. I do not have a copy of the North Caucasic Etymological Dictionary, so I was unable to fill in the evidential gaps indicated by Bengtson's use of the word 'etcetera'. Perhaps because of this I often struggled to relate the modern terms to the proposed ancestral forms, many of which appeared perverse in view of the incomplete supporting data cited. There were also several cases where the modern Caucasic forms were closer to the Basque than to the reconstruction, when the reverse should be the case. I did wonder whether such reconstructions were in fact correct.

Proper linguists may prefer to skip this next heretical paragraph, but I don't *like* reconstructions. I don't trust them. Reconstructions are theories, even the best reconstructions, yet too often they are treated as hard fact. It sometimes seems as if conformity to theory is prized above common sense, particularly when the reconstruction proposed appears unlike any word ever found in the real world.

Of course, Caucasic specialists will argue that their languages are typologically exceptional, and I am no more in a position to deny that than to deny the logical possibility that phonetic combinations may once have occurred that do not occur at present. Nevertheless, I believe were are justified in being suspicious of anything that has an asterisk in front of it and no supporting evidence.

I should add that there are many cases where the Basque word reinforces the proposed Proto-Caucasic reconstruction, and where the Proto-Caucasic form explains variations within Basque. This is how it should be.

This is also what I meant by 'non-random similarities'.

One of the greatest benefits of deeper level groupings such as Macro-Caucasic and Dene-Caucasic is that they allow 'out-group comparison' for their constituent parts, which is the best corrective to the excesses that can arise whenever reconstructions are 'internal-only'. Nowhere will this benefit be more apparent than in the prehistory of Basque.

My most serious reservations concern the question of time depth. First of all, I would dispute Swadesh's figure of 12-36% matches after 2,500-3,000 years. At a retention rate of 82% per thousand years, there will be 37.07% matches after 5,000 years and 13.74% matches after 10,000 years.

Secondly, neither the 29% nor the 41% matches have the meaning claimed for them because they are not based on comparisons between like and like. Equations derived from a particular set of criteria can only be applied where those criteria have been strictly adhered to. Every time the parameters are changed, the mathematical implications change too. They cannot be mixed and matched. For instance, because Basque is a single language spoken today, it can only be compared *mathematically* with another single language spoken today. Proto-Caucasic can, mathematically, only be compared to the version of Basque contemporaneous with it (now lost). Where examples from different dialects of Basque are included it is necessary to make a separate series of comparisons for each separate dialect of Basque and take an average. Only when such calculations have been made for every member of Caucasic (not that Swadesh lists for Caucasic are available in print!) can we even begin to try and date the relationship.

The *Cambridge Encylcopedia of Language* includes a table based on a complete series of pairwise comparisons within Caucasic carried out by J.C.Catford; this suggests an 11,000 year separation between East and West Caucasic. Although Bengtson's comparisons are mostly between Basque and East Caucasic, he maintains that the two branches of Caucasic are closer to each other than either is to Basque. Catford's figures suggest that Basque is separated from Caucasic (and, by implication, Burushaski) by more than 11,000 years.

Even if we could devise some way of quantifying how many matches language *x* will have to a family with *y* number of languages after *z* years, this would still not produce Bengtson's 29% and 41% because he does not confine himself to Proto-Caucasic forms. Comparisons would have to be *only* to Proto-Caucasic, *only* to Proto-East Caucasic, *only* to Proto-Circassian etc.

Finally, the figures for 'non-exact' matches are not what they seem either. A match between, say Basque 'elbow' and Caucasic 'knee' should not mathematically be treated as a single comparison. By implication, it is derived from two comparisons, between Basque 'knee' and Caucasic 'knee', and between Basque 'elbow' and Caucasic 'elbow'. A shared form involving words for, say, 'tree', 'wood' and 'stick' in three different languages is the result of three comparisons, not one. Every time more than one meaning is included this increases the total number of meanings used, and this reduces the percentages of meanings shared.

Obviously this has no bearing on the validity (or otherwise) of the comparisons made, or on how the frequency of matches found compares to matches found with any other language. These are solid in their own right. Nor does it make the connection any more or less 'obvious'. What it does do is challenge our perception of what 'obvious' actually means in terms of time-depth.

If the 'Splendid Isolation' of Basque is a glass ceiling that needs to be broken, so is the '5,000 Years Barrier'. Perhaps because of the Indo-European tradition, this particular time lapse has acquired an almost mythic status. Thus, because Indo-European is 'obvious' and has been dated to 5,000 years, everything 'obvious' must also be 5,000 years old. Any comparison linking Indo-European to some other 'obvious' family justifies doubling it to 10,000 — but beyond *that* there be dragons!

The problems of absolute dating are too many to go into here, but my suspicion is that many genetic relationships are older than is generally thought, even those considered 'obvious'. For instance, Joseph Greenberg's first pan-African classification consisted of sixteen separate groupings "in the class of the obvious". These included such diverse entities as Niger-Congo, and Afrasian. I seem to recall an early edition of *Mother Tongue* (the newsletter) in which Hal Fleming dated Proto-Afrasian to 24,000 years (+/- 8,000); in other words, this 'obvious' family may be between 16 and 32 thousand years old! As I said, the difficulties attached to absolute dating may yet prove insurmountable, but we should nevertheless prepare ourselves for the possibility that many of our guesses as to time depth, even in 'obvious' cases, may turn out to be seriously low. I am certain that the separation between Basque and the other Macro-Caucasic languages is *many* times greater than 2,500-3,000 years.

I am of course aware that there are other dates attributed to Indo-European and other benchmark figures in circulation. The differences between them are not important. My objection is to the concept of the Magic Number.

The one figure we have not yet looked at is zero, as in 'zero evidence'. The definition of 'evidence' may be debated, but 'zero' is fairly uncontroversial. If I had been unable to find a single point of apparent similarity between Basque and any or all other languages, that would be 'zero evidence'. Trask's use of the phrase is relation to Basque and Macro-Caucasic is not appropriate.

On the contrary, I believe that the time has come for Vasconists to put their energies into the job of refining Macro-Caucasic, instead of continuing to deny it. In so doing they may help themselves as much as they will help everyone else.

Response to Discussants

By John D. Bengtson

First, thanks to every discussant for responding, and providing a wide range of perspectives on the problem of the genetic classification of Basque.

Blažek: Not surprisingly, my friend and co-author (Blažek & Bengtson 1995; Bengtson & Blažek 2000), generally agrees with me on the issue of the genetic classification of Basque, while, naturally, disagreeing on some details. Blažek is generally more likely than I to accept the Romance origin of some words, e.g., Basque *ezagutu* 'to know', *hondar* 'sand', *buztan* 'tail' (though in the last case Blažek also provides an alternative Dene-Caucasian derivation!). Most of the words in question are rather peripheral, and do not touch the core of the most solid Basque-Caucasian comparisons (e.g., Basque *su* 'fire', *ni* 'I', *hortz* 'tooth' = Lak *ċu* 'fire', *na* 'I', *k:arč:i* 'tooth'), where there has never been any question of Latin/Romance origin.

On the question of the relative depth of the Basque-Caucasian relationship, Blažek compares the present day languages Basque and Lak, finds eight common cognates within the 100-word list, and estimates the time-depth of the Basque-Caucasian divergence to be comparable to that of Finnish (Uralic) and German (Indo-European), thus a macro-family depth rather than the "old family" depth proposed by me. After reading Blažek's paper, I conducted another experiment, comparing three East Caucasian languages with Basque, using the 100-word lists found in the "North Caucasian Database" diskette by Nikolayev & Starostin. (See Table 1.) Of course, the precise number of matches is, to some extent, subjective, and others might find fewer, or more, matches. Until the comparative phonology and lexicon of Macro-Caucasian are more fully worked out, some of the comparisons will remain uncertain. My critics might be surprised to know that I have purposely eliminated some very tempting 'look-alikes'. For example, the comparison of Basque *k(h)e* 'smoke' with Avar *k:uj* 'smoke' (and Bezhta *ko,* etc.) is beguiling, but Caucasian historical phonology shows that the Proto-Caucasian form contained a nasal (***k̲wɨ̆nħ**V NCED 738), and since Basque always faithfully preserves stem-final nasals (as reconstructed for Proto-Caucasian), this tantalizing 'look-alike' was eliminated from the table.

Basque *oin* 'foot' = PEC ***ʔĭnɢwV̌** 'heel' (No. 27, above)
Basque *oihan* 'forest, woods' = PNC ***fănV** 'mountain, hill' (No. 103, above)
Basque *akain* 'tick' = PNC ***q̱ā̆nʔV** 'louse' (NCED 911; see also MT II, p. 104)
Basque *muin* 'brain, marrow' = PEC ***mā̆fnɨ̄** 'brain' (NCED 797; MT II, p. 106)
Basque *zain* 'nerve, blood vessel' = PEC ***s̱ē̆ħmV** 'muscle, vein; intestine' (NCED 959)

Another such look-alike pair was Basque *gau* 'night' and Lak *x:u* 'night', but since the latter derives from a proto-form with laterals (PNC ***ʔlēƛă** NCED 216), and is *less* similar to Basque as we go back in time, the comparison was not included.

Even taking into account a margin of error, the results of Table 1 suggest that the time-depth of the Basque-Caucasian dispersal might be somewhat less than what Blažek

Table 1: Comparison of Four Modern Languages

(Vasco-centric: i.e., only Chechen, Avar, and Lak words
are cited that have a probable cognate in Basque.)

(gloss)	**Basque**	**Chechen**	**Avar**	**Lak**
'breast':	*ugatz*	-	-	*qazam*
'dog':	*(h)or*	*pħu* 'male dog'	*hoj*	-
'dry':	*agor ~ igar*	d-*eqa*	-*q̇:ʷara*-b	*q̇awq̇-*
'ear':	*belarri*	*lerg*	-	-
'earth':	*lur*	*moχk*[1]	*raƛ̣̆:*	*luχč:i*
'eye':	*begi*	*bʕärg*	*ber*	*ja*[1]
'far':	*(h)urrun*	-	-	*arχ(:)-*
'fire':	*su*	*c̣e*	*c̣a*	*c̣u*
'fly' (v):	*hegaz* egin	-	-	li̯-*χ:a-*
'head':	*buru*	-	*beṭer*	*baḳ*[1]
'horn':	*adar*	*kur*[1]	*ƛ̆:ar*	-
'I':	*ni*	-	-	*na*
'many':	*asko*	-	-	*č̣a̰*-u
'not':	*ez ~ ze*	*ca*	he-*č̣o*	-
'sleep':	*lo* egin	*nab*[1]	*ƛ̣̆:i-ž-*	-
'small':	*tipi*	-	*hiṭína*-b	-
'star':	*izar*	-	*c̣:ʷa*	*c̣u*-ku (pl. *c̣ur*-t:i)
'this':	*hau(r)*	*hara*	*he*-b, *ha*-b	-
'thou':	*hi*	*ħo*	-	-
'tongue':	*mihi*	*mott*	*mac̣:*	*maz*
'tooth':	*hortz*	-	*(gožó* 'fang, canine tooth')	*k:arč:i*
'two':	*bi-(ga)*	-	*ḳi-go*	*ḳi-a*
'what':	*ze-r*	*stē-*	*s:un-* (oblique)	*s:a-* (obl.)
Number of matches:		13	15	15

---

1. In some cases there is little or no superficial resemblance between the words. It is only comparative phonology that allows us to identify some cognates. For example, the three Caucasian words for 'earth' above are derived from PNC ***lhĕmŁ̱wɨ̆** 'earth', itself closer in form to Basque *lur* than is any of the modern Caucasian words. For 'ear, eye, tongue', see Bengtson 1999.

estimates, i.e., an 'old family' taxonomic depth, comparable more to that of Indo-European or Sino-Tibetan than to Nostratic. Note also that while each of the Caucasian languages cited here has a similar range of matches (13 ~ 15) with Basque, the matches are mostly with *different* words in each language, except for the six meanings ('dry, earth, ear, eye, fire, tongue, what') where all four languages have preserved the same lexeme.

Schuhmacher: My own research has led me to agree with Vennemann, among others, that a language family related to Basque once dominated a large part of Europe (see, e.g., Bengtson 1991). Schuhmacher's note here raises interesting questions, and we hope he will discuss them in more detail in the future.

Sidwell: In *Mother Tongue* IV, Paul Sidwell compared 110 Proto-Ainu words with words from a wide array of Austroasiatic languages, and found 61 apparent matches. Sidwell concluded that "the evidence for a genetic relationship between Proto-Ainu and Mon-Khmer is real, and deserves further consideration." (By the way, I agree with Sidwell on this point.) In the study being discussed here, over 200 basic Basque words are compared with Caucasian words, and 106 apparent matches have been presented. In the latter case, however, Sidwell finds himself agreeing with Larry Trask that there is "zero evidence" for a genetic connection. Why the different conclusion? In both studies, essentially the same method was used, and comparable results were obtained.

In fact, in the Basque-Caucasian case the constraints were more stringent: whenever possible, etymologies common to the whole Caucasian family (PNC) or to one of its major subgroups (PEC, PWC) were cited, with reconstructions when available. Comparisons that were severely localized, such as the Tsakhur word for 'tail', were clearly marked as such, and relatively infrequent. In Sidwell's Ainu-Mon-Khmer study, one could not tell whether a word was typical of Austroasiatic or restricted to one or two languages. In addition, my Basque-Caucasian study went a step farther than Sidwell's in showing that a number of phonological correspondences have been found between Basque and Caucasian. On the face of the evidence alone, I am stymied as to why Sidwell in the first case finds the evidence for genetic connection "real," while in the second case it is "zero." I can only attribute it to his admiration for Trask.

Sidwell seems to think that one has to be "starry-eyed" to believe that long-range comparison is possible. I say there is nothing "starry-eyed" about it: we have scientific methods, and if we apply them carefully, we will have results (Starostin 1992). We have applied these methods to the Ainu-Austric case and the Basque-Caucasian case, and we conclude that in both cases genetic connection is the most probable explanation for the similarities. That is no more starry-eyed than the method of the paleo-anthropologist, who identifies bones as belonging to a particular species, or of the astronomer, who infers the existence of a planet by means of indirect evidence. Sidwell seems to be caught up in the paralyzing ideas recently publicized by R.L. Trask and R.M.W. Dixon, et al., claiming that long-range comparison is impossible. In my opinion, these ideas are unscientific, and lead to regression rather than progress in the science of historical linguistics.

Trask: Note that many of the Basque words were already discussed in *Mother Tongue* I (q.v.), and it is clear that Trask and I still disagree on many etymological questions. See my article in MT I (Bengtson 1995) for discussion of what I regard as errors of method by Trask and some other Vasconists. Some responses to Trask's comments on individual Basque words:

Basque *bul(h)ar* 'breast, chest': Trask ignores the most conservative form, *bulhar,* since he finds no etymological significance to Basque /h/ (see below). Trask argues against the primacy of *bul(h)ar* (vs. *burar* ~ *budar*), because in Zuberoan "original /u/ was fronted in that dialect … though not before the tap /r/." But Zuberoan has /bürü/ or /büü/ 'head' for common Basque *buru* /buru/.

Basque *(h)or* 'dog': Trask claims that the /h/ is secondary, since it is only found in Zuberoan, which is "the dialect which routinely adds aspiration to all the monosyllables which can bear it." As I showed in MT I (pp. 85-86), Zuberoan in fact maintains a contrast between aspirated and unaspirated monosyllables:

har 'worm' vs. *ar* 'male'
hots 'come on' vs. *ots* 'male (animal)'
hütz 'fart' vs. *ütz* 'to leave'

I think it more likely that the other Basque dialects have lost the aspiration in this word (and *oin* 'foot', which is *hün* in Zuberoan).

Basque *belarri* ~ *beharri* ~ *begarri* 'ear': the variant *beharri* was probably influenced by *beha-* 'to listen', and *begarri* [beγar̄i] is the result of voicing assimilation of *beharri.*

Basque *adar* 'horn' ~ Old Irish *adarc* 'horn': since the latter has no Indo-European etymology, it is probably borrowed from Vasconic **adar-ko* 'little horn'.

Basque *hanka* 'leg': see under Basque *zeru,* below.

Basque *medar* ~ *mear* ~ *mehar* 'narrow': Trask's interpretation is the reverse of mine. It seems more likely to me that *medar* is the original, and *me(h)ar* is secondary, by contamination with the originally unrelated *mehe* 'thin, slender'. There is no evidence that /d/ has ever been "inserted to break up the hiatus"!

Basque *zeru* 'sky': Trask is quite emphatic in his denial of a Caucasian connection: "No, no chance." José Ignacio Hualde also told me (in a letter) that the Latin derivation of *zeru* was "obvious" to him. Two things still trouble me about the word. First, all early Latin loanwords preserve the old velar, not the Romance palatalized affricate or fricative: Latin *certu* > Basque *gertu* 'certain'; Latin *cella* > *gela* 'room'; Latin *ceresea* > *gerezi* 'cherry', etc. (Trask 1997: 170). Latin *caelu* "should have" become **geru.* The other is the Roncalese form, *zeuru* ~ *zeuri.* Why would a simple disyllabic word acquire another syllable? (Roncalese *deuri* 'money' comes from a tri- or quadrisyllabic original.) So far, it seems that only Blažek agrees with me. But even if Trask, et al., are right about this, it would be no great loss to the evidence here. Much the same could be said about Basque *hanka* 'leg, haunch'.

Basque *el(h)ur* 'snow': Trask cites only *elur* ~ *erur* ~ *edur,* ignoring the /lh/ heard in the northern dialects (Basse Navarre, Lapurdi, Zuberoa). See below for more discussion of /lh/.

Basque *listu* ‘spit’: Trask provides an even more satisfactory comparandum here: “Navarrese … *lits* ‘small spot of liquid on the ground’ … originally … ‘gob of spit on the ground’.” (Compared with Tindi *λac:i* ‘saliva’, etc.)

Basque *makila* ‘stick’: “everybody agrees that [it] derives from Latin *bacilla … ”* But dialectal *maket* surely can’t. A merger by contamination is also possible.

Basque *mihi* ‘tongue’: Trask and I continue to disagree, he claiming that “the Pre-Basque word was _*bini_ and nothing else … can account for all the numerous regional variants.” Not quite, Professor Trask. It cannot account for the strong fricative recorded by Nicole Moutard in northern (Basse Navarre) Basque, as in [mih(ç)ja] ‘the tongue’, corresponding to literary *mihia.* Trask’s derivation requires _*n > h >_ [ç], clearly impossible. An extensive discussion of my reconstruction of Basque _*mixi_ ‘tongue’ is found in my article “’Eye, Ear, Tongue’ in Basque and East Caucasian” (Bengtson 1999). I urge the reader to carefully consider that evidence before deciding that _*mixi_ is incorrect.

Basque *emakume* and *emazte* ‘woman’: again, Trask and I continue to disagree. On the semantic side, words for ‘girl’ rarely if ever change to words for ‘woman’. In European languages, words for ‘girl’ (and ‘boy’) are highly variable, even within dialects of the same language (English *girl, lass, maid(en),* etc.; Swedish *flicka, gräbba, jänta, tös, tjej,* etc.), while words for ‘woman’ (and ‘man’) are much more stable. His last note, that Basque *eme* ‘female’ is “transparently” a borrowing from Romance is certainly debatable. Even Michel Morvan, who agrees with me on little else, thinks that *eme* is ancient and native. *Eme* is too deeply embedded in the lexical structure of Basque to have been adopted in recent centuries from a late Romance source. Several of Trask’s etymological errors are of this anachronistic or historically improbable nature (see Bengtson 1995, p. 89). It seems far more likely that both *emakume* and *emazte* are the ancient words for ‘woman/wife’, both found not only in Caucasian, but also in Yeniseian (e.g., Ket *qi·m* ‘woman’, corresponding to Basque *–kume,* Archi *χom*, etc.) and Na-Dene (e.g., Haida *ǯáadaa* ‘woman, girl’, corresponding to Basque *–zte,* Chechen *stē*, etc.).

Basque *har* ‘worm’: see Blažek’s discussion of this word.

Trask’s comments on Appendix A: “there can be no justification for regarding Basque /lh/ as a segment different from /l/. … even in the aspirating dialects, there is considerable variation …” But in any *one* aspirating dialect, there is a clear distinction between words that *always* have /l/ and others that *always* have /lh/. For example, in the northeastern Zuberoan dialect described by Larrasquet, *belhar* ‘hay’, *bilho* ‘hair’, and *olho* ‘oats’ always have /lh/, but *bele* ‘raven’, *bilaiz* ‘to undress’, and *xilo* [šilo] ‘hole’ always have /l/. This is a clear phonemic contrast, which Trask is at a loss to explain. I have provided an explanation with clear external cognates that explain the origin of the contrast.

Trask’s comments on Appendix B: “invoking ‘assimilation, dissimilation or metathesis’ … effectively allows you to match any two liquids at all.” Yes, the way we do with Spanish *milagro* and Latin *miraculu.*

“*merdar is indefensible: the etymon of Basque *mehar ~ medar* is _*bene + *-ar.”_ (And what is _*-ar?)_ As explained above, the association with *mehe* is probably secondary.

In his comment to the ‘who’ comparison, Trask again laments that “the numerous comparisons of Basque with language L *never* shed light on our problems.” In Appendix

Table 2: Comparison of Dene-Caucasian Class Morphemes

(In Sino-Tibetan and Basque, the morphemes became lexicalized (fossilized);
in Proto-Caucasian they were markers of a living 4-class system,
with eventual fossilized remnants in modern Caucasian languages.)

SINO-TIBETAN	CAUCASIAN			BASQUE
	per Schulze-Fürhoff	class	per Diakonoff & Starostin	
m-	**w-**	*1 sg.*	**u̯-**	**u-/o-**
r-	**y-/r-**	*2 sg.*	**j- (= y-)**	**i-/e-**
b-	**b-**	*3 sg., 1,2 pl.*	**w-**	**bi-/be-**
d-	**d-**	*4 sg., 3,4 pl.*	**r-**	**ar-**

Some examples of fossilized prefixes in Caucasian languages: Rutul *u-xun* and Tsakhur *wu-xun* 'belly', Tindi *b̄-eƛ̄:u* 'stomach, rennet', Hunzib *b-eƛ̄* 'rennet', Ubykh *b-La* 'eye', Circassian *p-śa* 'neck' (with PNC ***w-**); Lezgi *ru-fun* 'belly', *r-uš* 'daughter' (with PNC ***r-**); Ubykh *t-χamə́* 'skin, fur', *č́a* 'milk' < ***T-š́ə** (with PNC ***T-** or ***d-**); some Avar-Andian languages also show traces of a prefix ***m-**, reminiscent of Sino-Tibetan ***m-**: Karata *m-eƛ̄:u* 'stomach', *mi-gaž* 'beard'; Avar *ma-xá* 'abomasum', *mi-λír* 'wing', *me-héd* 'brisket', *me-géž* 'beard'. (Cf. Tibetan *m-čhin* 'liver', *m-gul* ~ *m-gur* 'throat, neck', *m-čhu* 'lip, beak', etc.)

For more detailed discussions of the postulated Dene-Caucasian class prefixes, see Bengtson (1995, 1997, 1998).

A., I provided an explanation of the Basque contrast between /l/ and /lh/, which Trask apparently ignores as insignificant. In Appendix C., I repeat the old comparison of Basque *negar* ~ *nigar* 'tears' and Caucasian ***nĕwqŭ** 'tears, pus', showing that the Basque word is both lexically and morphologically comparable with Dargwa (Akushi, Chirag) *nerʁ* 'tear', itself a fossilized plural derived from earlier ***neʁ^{w}-r.** Time after time I have discovered solutions to Basque etymological problems in external comparison with Caucasian, and other Dene-Caucasian languages, but Trask seems to find none of them significant. One wonders what would be.

Vajda: I read Professor Vajda's article with great interest. With some reservations, I agree with his call for a demonstration that Basque and Caucasian (and other Dene-Caucasian languages) are characterized by a "*complex of shared structural features* which represent obvious genetic inheritance." One reservation derives from the fact that in long-range comparison, we have found that morphological structures can change radically, and adapt to those of 'unrelated' languages by areal contact. Thus, the earliest recoverable forms of Sino-Tibetan have already lost much of the postulated morphology of Dene-Caucasian, and Na-Dene morphology has probably been deeply influenced by contact with Amerind and Eskimo-Aleut languages. Even in spite of this, I have identified a 'complex of shared structural features' that is common to Basque, Caucasian, and Sino-Tibetan. In *General Linguistics* (Bengtson 1998: 37) I presented a table comparing the Sino-Tibetan nominal prefixes with Caucasian class markers and the fossilized class markers of Basque. (See Table 2.) As noted there, this scheme is so far provisional, but it is a first step in outlining a 'complex of shared structural features' called for by Vajda.

Vajda is the only discussant to bring up an important issue: the validity of the (North) Caucasian language itself, which is not universally accepted by historical linguists. I have assumed the validity, based on the work completed some six or seven decades ago, by, for example, N.S. Trubetskoy (1930) and Georges Dumézil (1933), outlining the lexical and morphological affinity of the West (Abkhazo-Adygean) and East (Nakh-Daghestan) Caucasian families. This work is now supplemented by the recent work of Nikolayev & Starostin (1994) and Vyacheslav Chirikba (1996), describing in detail the lexical and phonological unity of the same two families. If the validity of this family is still doubted by some Caucasian experts, I have to ask – on what basis?

Vajda also calls attention to the list of "puzzles" compiled by Trask (1996: 115-116), and notes that "Basque will remain an isolate … in the minds of most future linguists until someone succeeds in providing plausible answers to at least some of Trask's questions." I can only agree, and call upon my colleagues (as I have repeatedly before) to work with me, or independently if they prefer, on these questions. I have already proposed an answer to part of Trask's first question, about "an extraordinarily large proportion of [Basque] lexical items beginning with a vowel …" I have suggested that many of these words may be the result of fossilized class markers, as seen in Table 2, for example:

Basque *uzki* 'anus' (*u-zki:* cf. Tibetan *skyi-ša* id.)
odol 'blood' (*o-dol:* cf. Na-Dene ***deɬ** 'blood')
izar 'star' (*i-zar:* cf. Caucasian ***ʒwhărī** 'star')

izen ~ *uzen* ‘name’ (*i-zen* ~ *u-zen:* cf. Burushaski *sén-as* ‘said, named’)
el(h)ur ‘snow’ (*e-l(h)u-r:* cf. Caucasian ***λĭwV*** ‘snow’)
esne ‘milk’ (*e-sne:* cf. Caucasian ***š̱änʔu*** ‘milk’: MT-II, p. 105)
egur ‘firewood’ (*e-gur:* cf. Caucasian ***gōrV*** ‘pole’), etc.

Trask’s main objection to this is the fact that “there are no prefixes in Basque,” which is irrelevant, since I am postulating *fossilized* prefixes. (See, e.g., Bengtson 1997.)

But Vajda is right: much more remains to be done before the questions raised by Dene-Caucasian are answered to the satisfaction of most historical linguists. I will continue to work on this deficit, and I call on others, hopefully young linguists, to help.

Whitehouse: Whitehouse discusses an important point: the process of elimination that precedes a hypothesis of classification. I have never publicly discussed the process that led to my deep interest in Basque and Dene-Caucasian, which was as follows: in the early 1980’s I began an intensive multilateral comparison of all human languages, to see for myself whether there was any evidence for a Proto-Human language, or for a deeper language classification than was commonly accepted. I was already acquainted with the evidence for Indo-European (with assistance from Eric Hamp, while I was still a teenager!), AfroAsiatic, Bantu, Austronesian, and some other accepted families. I plunged into wide-ranging studies of African, American, Asian, Oceanic, and American languages, in the process beginning to identify parts of some global etymologies, as well as some of the diagnostic lexemes that distinguish larger genetic groups. Still unaware of Greenberg’s Amerind hypothesis, I found that there were deep connections between North and South American languages, as well as with the Old World languages. After this I met and corresponded with some colleagues who had an important influence on my thinking: Roger Wescott, Vitaly Shevoroshkin, and Merritt Ruhlen, all of whom had done, or were doing, some of the the same kinds of research I was doing. Shevoroshkin introduced me to the Russian Nostratic school and provided me copies of their work, which again stimulated renewed research.

It was after all this that I began to work with Basque, and soon found that it did not belong in Nostratic, Afro-Asiatic, nor any other family I was acquainted with. But there was a family some of the Russian Nostraticists had (re-)discovered, which they called “Sino-Caucasian.” After applying multilateral comparison of Basque with Nostratic and Sino-Caucasian, I found that it, along with Burushaski, belonged with Sino-Caucasian (now more commonly known as Dene-Caucasian). Years later, Joseph Greenberg told me that he had independently come to the same result as I had, though he never published the conclusion until recently, stating his agreement with Dene-Caucasian. Now Paul Whitehouse has independently come to the same conclusion. This is a far cry from what some critics have imagined: that we have invented a “dustbin” macro-family, into which we sweep all troublesome isolates! If this were so, would we not have tried to “force” Ainu into it? (Ainu’s a darn sight closer [geographically] to Sino-Tibetan than Na-Dene is.) To the contrary, I have never suggested the inclusion of Ainu in Dene-Caucasian, and have consistently connected it with Austric (Bengtson 1992; Bengtson & Blažek 2000). As I stated in another paper (Bengtson 1998: 46), “**evidence selects the membership,** i.e., diagnostic lexical and morphological elements indicate which family or macro-family a language belongs to.”

References:

Bengtson, J.D. 1991. "On Dene-Caucasian Substratum in Europe." In *Dene-Sino-Caucasian Languages,* ed. by V. Shevoroshkin, p. 171. Bochum: Brockmeyer.

---------- 1992. "A Case for the Austric Affiliation of Ainu." In Shevoroshkin (Ed.) 1992, p. 364.

---------- 1995. "Basque: an orphan forever? A response to Trask." *Mother Tongue* I: 84-103.

----------1997. "Basque and the Other Dene-Caucasic Languages." In *The Twenty-Third LACUS Forum 1996,* ed. by A.K. Melby, pp. 137-148. Chapel Hill, NC: LACUS.

---------- 1998. "Caucasian and Sino-Tibetan: A Hypothesis of S.A. Starostin." *General Linguistics* 36(1/2): 33-49.

---------- 1999. "'Eye, Ear, Tongue' in Basque and East Caucasian." In *From Neanderthal to Easter Island (Tribute to W.W. Schuhmacher),* ed. by N.A. Kirk and P.J. Sidwell, pp. 3-10. Melbourne: Association for the History of Language.

Bengtson, J.D. and V. Blažek. 2000. "Lexical Parallels between Ainu and Austric, and Their Implications." *Archiv Orientální* 68(2): 237-258.

Blažek, V., and J.D. Bengtson.1995. "Lexica Dene-Caucasica." *Central Asiatic Journal* 39: 11-50, 161-164.

Dumézil, Georges. 1933. *Introduction à la grammaire comparée des langues caucasiques du Nord.* Paris.

Shevoroshkin, V. (Ed.). 1992. *Nostratic, Dene-Caucasian, Austric and Amerind.* Bochum: Brockmeyer.

Starostin, S.A. 1992. "Methodology of Long-Range Comparison." In Shevoroshkin (Ed.) 1992, pp. 75-79.

Trask, R.L. 1996. "Response to Starostin." *Mother Tongue* 2: 111-117.

---------- 1997. *The History of Basque.* London/New York: Routledge.

Trubetskoy [Trubetzkoi], N.S. 1930. "Nordkaukasische Wortgleichungen." *Wiener Zeitschrift für die Kunde des Morgenlandes* 37/2. Vienna.

An Enquiry about Sumerian Words

By Liny Asoka Srinivasan
Monmouth University
West Long Branch, New Jersey, U.S.A.

This author is fascinated by the similarity of Sumerian to so many different languages, as put forward by various experts in *Mother Tongue* (Issue III, 1997), as well as encouraged to write this enquiry about Sumerian words in *Desi* vocabulary. The joint paper of this author with Cyrus Gordon[1] and a few follow-up articles in Bengali [2] of this author, have been able to convince most of the Bengali scholars that a large number of *Desi* [Skt. *deśi* 'belonging to the country'] words of unknown origin in Bengali are from a Canaanite source. Since then Malati J. Shendge (a linguist and Indologist), in her recently published book,[3] has shown about four hundred cognates between Akkadian and Sanskrit. According to Sanskrit and Prakrit grammarians, *Desi* words are non-Sanskrit, and they are found both in Sanskrit and Prakrit from time immemorial.[4] In brief, the non-Sanskrit origin of *Desi* words was recognized by the scholars of the classical era; and modern scholars also believe that they are not related to any Indo-European / Indo-Aryan source. It was inferred from the meaning of the word *Desi* that all such words of unknown origin in all major North Indic tongues (NIA), particularly in the spoken and colloquial dialects, have come from distant tribes who inhabited the land before the immigration of the Aryans. Recently that idea of deriving the *Desi* words from native tribal sources is changing slowly, as many *Desi* words seem to bear direct relationship with words from Semitic as well as non-Semitic ancient Near Eastern languages, such as Sumerian. It is worth mentioning here that at least two Sumerian words, namely *paraśu* 'axe' and *apsu* 'water \ watery', were found in the Rigveda by the eminent Indian linguist S. K. Chatterjee during the fifties[5]. The following few examples of *Desi* words and their similarity to Sumerian words are presented to Assyriologists, linguists, anthropologists, and others for their opinion.

Sumerian	**Desi Bengali (or as indicated)**
giš 'tree, wood'	gä$_{\text{Ç}}$ (for palatalization of s/š see below) 'tree'
giri 'foot'	gor̲ (r̲ = retroflex r) 'foot, heel' also in Prakrit
gi-gir 'chariot, wagon'	gär̲i 'any vehicle'
ge-en (phon. writing of gin) 'like,as'	jena (j = soft g) 'like, as'
ga-ar 'wall'	ghäro (g = gh, Prakrit) 'wall, rampart, enclosure, *Desi*/Prakrit ghar 'room, house, dwelling place, residence '
udu 'sheep'	udo 'sheep' (used idiomatically for 'stupid'), hudo (Prakrit.) 'sheep'
úr 'thigh,	úru 'thigh, upper part of the legs'

uru ‘city, town’	úr (old Bengali), úro (Prakrit) ‘city,town’
utu ‘the sun, the sun god’	itu ‘the sun, the sun worshipped as a god’; etu ‘the sun’ (cf. Hindi etowar ‘Sunday’)
uzu ‘priest’	ojhä ‘diviner, magic curer’
mi ‘woman’, determinative for ‘female’	meye, mäiä ‘woman, female’, mäi (Middle Bengali) ‘an elderly woman, mother’
mà-a ‘I’	mui, mai (Eastern dialects, Hindi) ‘I’
im ‘mud, clay’	ämä (Middle Bengali) ‘things made of baked or unburned clay’
i-lu ‘song, wail’	ulu ‘ceremonial and auspicious sound made by women in wedding/religious festivities’
dim ‘pillar’	thäm ‘pillar’
dug ‘jug, pot’	ḏhäk ‘large drum shaped container, drum’, dug-dugi ‘small drum as musical instrument
dur ‘bond’	dor ‘bond, bindings’
dá-ra ‘a band’	ḏaṟä, ḏaṟi ‘a band, rope’
bar ‘outside’	bär ‘outside’
bar, bara ‘seat or shrine of a god’	bärä (Middle Bengali) ‘enclosed spot of a deity’
pal ‘turn in office’	pälä ‘turn in office, time of one’s turn’
sar ‘whole, all’	särä ‘whole, all’
šá-an ‘clever’	seyanä , seana, seyän ‘clever’
rin-na ‘oven’	rännä ‘cooking, to cook (with aux.)’
nitah ‘male’	nitä ‘male worker’ (colloq.)
tum ‘abundance, plenty’	dhum ‘plentiful, abundant, ostentatious display of abundance’
lu-gal ‘king’	laulä (eastern Prakrit) ‘king’, raulä (Middle Bengali) ‘ king, royal’
lu ‘man, fellow’	loa (Old Bengali) ‘man, fellow’

The above and all examples in this paper are Bengali (West Bengal dialects), current or from Old and Middle Bengali literature.[6] Some of the words are found in Prakrit literature, and a few occur in neighboring dialects, such as Hindi. In brief these are dialects classified as Middle or Modern Indo-Aryan. None of them are found in Dravidian, with the exception of *uru* ‘town’, which then needs some explanation. It should be mentioned that the Brown University Department of Egyptology has recently published C. H. Gordon’s article “The Near Eastern Background of the Rigveda”. [7] Sumerian *uru* ‘town, determinative for town’ is similar to Dravidian *uru* (Telegu) and *ur* (Tamil) ‘town, settlement’, but the word is also present both in Prakrit and Old Bengali.

We cannot now determine the manner of dissemination of this word, because this word is very likely present in the Rigveda, the oldest text of India. Its hidden existence is detectable in the Rigveda in the form of topophorous royal / ethnic names such as *Uru-kṣaya, Uru-kakṣa.* These are royal titles and *rishi* (sage) names, but their use in the plural as *Urukshayas* for a people or group of *rishis,* and the fact that according to the traditional commentators *Uru-kshaya* of the Rigveda was a place, suggests that perhaps the word *uru* was prefixed to indicate that *Kshaya* was a town. That also explains why sometimes, in the Rigveda as well as in later epic-puranic literature, *Kaksha* appears as a titular name for kings, *rishis,* or a people without the prefix *uru* (where such names in general are topophorous).

It is also interesting that, besides the substantives, many *Desi* verbal roots also seem to be closely related to Sumerian verb stems. According to Allan R. Bomhard "the Sumerian root was generally monosyllabic; *CV, VC,* and most often *CVC.* There was no distinction between verbal roots and nominal roots."[8] This is also true for *Desi* words, except for some words that have been Sanskritized and extended. In Bengali, the verbal root is also used as the imperative mood of a verb: e.g. *khä* 'eat' is both a verbal root and a command or strong request to eat. In other words, in Bengali the simplest form of verbal stems and the inferior imperatives are the same. Unlike Sanskrit, use of auxiliary verbs is an important feature of Bengali, which often keeps the main verbal roots uninflected and provides the flexibility for a stem to be used either as a noun or as a verb. These characteristics have been studied in great detail by linguists, and particularly by S.K. Chatterjee in his book *Origin and Development of the Bengali Language;* and the following paragraph is a brief summary of his explanations for such peculiar non-Indo-Aryan Bengali verbs. According to Chatterjee, there are 1500 roots in Bengali, but 1300 of them cannot be traced to OIA or Sanskrit, and so they were conjectured by him to have evolved from original root plus affix forms. He defined them as bare roots, and, in his own words, "what was originally a combination of root plus affix has commonly been reduced to the bare roots through phonetical decay." He also called it a root noun. "What at first looks like the root unmodified by any inflection features as a verbal noun and also as the inferior imperative" (note. 5. vol. II. p. 895). However, at present these explanations are inadequate in view of the fact that many such eroded or bare roots or so called root-nouns seem to have exact counterparts in words of various ancient Near Eastern languages. The following examples are comparisons with Sumerian verbal forms.

Sumerian	**Bengali**
qar 'to take away'	käṟ 'to take away by force'
gar ' to do, make, establish'	gaṟä 'to do, make, build, establish (something creative like making dolls or utensils from clay, building a house, or things glorious such as establishing a city or organization)
gúr 'to bend, bow down'	gaṟ 'to bend down to show

	respect’
gur ‘to grind, to rub down’	gur̲ä, gur̲o ‘to grind, to powder down’
gúr ‘to twist, turn, roll up’	gar̲ä ‘to roll, roll on the bed or floor’
gu-tab ‘to twist yarn’	gutä ‘to twist yarn, to roll up things like carpet, paper ‘
bad ‘to separate, to divide, to part’	bäd ‘ to separate, discard, discount, subtract’
par ‘to go or pass by, to go past’	pär ‘ to pass by, to go past, to go across’
pá(d) ‘to call. to speak’	pär´ (r = retroflex ḍ) e.g., däk pär ‘call’, hãk- pär ‘shout’
pad ‘to tear out, remove’	pär ‘to tear out (fruits/flowers), tear down, remove husks from paddy’
men ‘to agree, allow’	mänä ‘to agree, to concede, to abide’
nad / nu (-d) ‘to lie down, to sleep’	nid ‘to sleep’
nir-nir ‘winnow’	nar̲, nar̲-nar̲ ‘to move, shake’; nar̲-nar̲e (adjectival)
sar ‘to write’	sarä ‘to write' as in kalam-sarä ‘to write with a pen’
šár ‘to bring together, to organize’	sär-sär, säri, ‘to collect or organize things’
sar ‘to drive away, to drive forward’	sarä ‘to move things, to remove, to drive away’
si ‘to lay’	šo ‘to lay down, to lie down, to sleep’
til, til-til ‘to pick off, to pluck’	tolä ‘to pick off, to pluck’
te-(gá) ‘to attain, to reach, to touch upon, to meet’	ṭhekä ‘to reach, to touch upon, to meet’
tag ‘to spoil’	t̲ak as a noun means ‘sour’ but as a verb means ‘to spoil’
tar(r) ‘to cut’; dar ‘to split’	thor̲ / thor̲ä ‘to chop, to rend’
tar(-r) ‘, to break’	tor̲, tor̲ä ‘to break’
tuk ‘to have done’	thäk ‘had done’

Words with an unaspirated sound like /g/ sometimes interchange with words with /g/ between literary or good Bengali and its dialectical forms. For example, Sumerian *gur* ‘to return, come back ‘ is found to be the same as *gur* ‘to return, come back’ in the Manbhum district of West Bengal. However, in “good” (standard) Bengali it is *ghur*. This also seems to happen with Sumerian words. For example:

Sumerian	**Bengali**
gur 'to run about'	ghurä 'to travel, travel around a place'
gur-gur 'to go hither and thither'	ghur-ghur, ghorä-ghuri 'to go here and there, to go about a place, to go around'
gur-gur 'to reel'	ghurä, ghoräna 'to reel'
gur 'to return, come back'	ghur (with aux.) 'to return, come back, give back'
guz 'to crush, to smite'	ghuc, ghucä 'end, destroy with force or violence'
gir / ger ' to enclose'	ghir / gherä 'to enclose, to surround'

The velar /k/ in Sanskrit words frequently interchanges with /kh/ in Pali, Prakrit and modern *Desi* words. Sumerian examples would be:

Sumerian	**Bengali**
ku 'to eat'	khä 'to eat'
kuš 'skin, hide, leather'	khosä 'skin (of vegetables), khos 'scab, skin desease' khuski 'dandruff, the dry, flaky skin of the scalp' khas-khase (idiomatic) 'dry like leather'

Some Canaanite words are found recently to have survived in *Desi* vocabulary being relegated to very specific use. This is also perhaps the case with some good Sumerian words. For example, Sumerian *tibir* 'palm', a blow with the palm of the hand' is well preserved in the words *thäbrä* and *thäbar* / *thäppar* 'patting / slapping with the palm'. This may indicate that *thäbä* 'palm' is a short form of original *tibir*. *Thäbä* is still used for the palm of the human hand, yet it is more commonly 'animal's paw'. Since systematic studies of *Desi* words have not been done, and only following the discovery of Canaanite words in Bengali, some dictionaries or collections of Bengali *Desi* words have appeared in the bookstalls of Calcutta, it is hard to judge how many different ways Sumerian words can be hidden in the vast range of so-called *Desi* vocabulary. First and foremost of them is the regular pattern of slight phonetic changes, which may be shown here in the form of alphabetic changes. They are detectable because their meanings are identical. For example, just as /s/ and /Š/ (sh) of Sanskrit words are often palatalized in the *Desi* forms; similarly /s, š/ of Canaanite and Sumerian words change to /c/ (ch as in China) and /ç/ (ch as in chubby), while /z/ could interchange with /j/ or /jh/. For example, *Desi* *çoot* 'to run' has been compared to Hebrew *šût* 'to go to and fro' (cf. note 1). The Sumerian examples would be: sir 'to tear' and *Desi* *çiṟ* / *çeṟ* 'to tear'. Sumerian *túg-mu-sir-ra* 'rags, tatters' and *Desi* *çeṟä* 'rags, tatters, torn or scrapings. This raises the question whether Sumerian *šu,* 'hand' as well as 'to do things with hands', could be compared to the *Desi* verb *ço, çu* 'to touch, feel with hands'? Ever since Chatterjee's book ODBL [see above], sanskritization of *Desi* words has been accepted as an invisible process that has been going on in India from an unknown past. In brief, sanskritization of *Desi* words, according to Chatterjee, means adaptation of words from various local vernaculars into Sanskrit by slight phonetical modifications or extensions at the end, and passing them off in written vernaculars as Sanskrit words, albeit the Near Eastern connections or sources of the *Desi* words were neither known nor even suspected. This can further complicate the Sumerian issue. For example, could *šur̠* 'tenticles, trunk (used like hands)', being a *Desi* term, be a crudely Sanskritized extension of Sumerian *šu* 'hand'? Thus the number of examples of possible Sumerian origin increases greatly if allowances are made for slight phonetical changes, changes of forms between verbal, nominal or adjectival and, consequently, slight shifting of the meaning as exemplified below.

Sumerian	**Desi Bengali (or as indicated)**
a-aga 'instruction, decree' a--ag 'to give comand to'	ãggä (ã = nasal) 'instruction, order, command of a superior or master'
aga 'crown, tiara, i.e., thing worn on top of the head'.	ägä 'top or tip of things like that of a tree, in Middle Bengali used for 'head, top'
ér 'water'	irä (Prakrit) 'water'
ig 'door'	ägal (poetic) 'door'
bui 'knowledge, learning'	bai 'book' (this unique *Desi* word once meant 'book of knowledge', like Pali and Sanskrit *śāstra* 'book of knowledge'
gun 'tribute, talent'	guna-gär ' to pay or compensate by money'
sir 'to be long'	cir 'very long time, whole or entire life, life long, forever'
kur 'mountain'	cur̠, cur̠ä 'peak of a mountain'
kar 'town'	gar̠ 'fortified town'
tag 'to weave'	ṯäku 'spindle of a loom'
ku, kua 'fish'	keut 'fisherman' (Middle Beng.), keor´ä 'a caste of fisherman'
kemu 'flour'	gam 'wheat'
sur 'to squeeze, to press out'	šur̠i 'the caste name of wine makers and sellers', i.e., 'those who used to press grapes'
sila 'measure of capacity'	sera 'a traditional unit of capacity'
iku 'unit of area'	ekar 'traditional unit of area'
izi 'fire'	ãc 'fire'
tur 'to be small in size, weight'; tur-ra 'small, childish'	thorä 'little, small, insignificant in amount, weight, quantity, importance'
guru 'a large measure of grain'	guru 'adjective for heavy weight'
gur 'hefty, thick'	guru 'heavy, serious, important'
ma 'boat'	mä-ît (Old Bengali) 'the front part of a boat'

ma-lah / má-lal-a 'boat man, sailors'	mälä, mallä 'the caste name for mariners, sailors, fishermen'
^{giš} má-šu (-a) 'ship, deep draught boat'	mäjhi 'a caste of boatmen'
mun 'salt'	nun 'salt'
na-ru-a 'stela'; Akkadian naru 'stone for monuments, boundary stones'	noṟä 'term for round and elongated grinding stone'
zal 'to flow'	jal 'water'
zi 'life'	ji , jio 'to live'

It has always been assumed that *Desi jal* has come from Sanskrit *jala-* 'water'. But it appears only in late Sanskrit, later than its common use in Prakrit texts . Could it be an example of a Sanskritized form? Sumerian *zalag* 'to cleanse, to purify' shows that the sense 'water' was perhaps imbedded in the Sumerian stem *zal,* because cleansing or purifying implies washing with water. Tri-consonantal words in Sumerian are very few, yet there are some *Desi* examples as follows:

Sumerian	**Desi Bengali (or as indicated)**
da-gal 'wide, broad'	dägar (poetic word) for 'large, big, broad'
za-lag 'shining, bright'	jhalak 'shining, flashing, bright'
su-húr 'braids, chignon'	cịkur poetic word for 'carefully done hair or hairdo'
nin-da 'bread, food in general'	rän-dhä 'cooked food'
nin-da 'council'	nindä 'criticism, saying bad things about a person'
nagar 'carpenter'	lagur 'wooden stuff, timber'
ka-lal 'honey-mouth', used for sweet words (cf. W. Hallo)	kallol 'sweet sounds, bird's songs'
kik-kin ' fine milled'	cikan 'very fine'
ki-kal-la 'hard and dry ground, barren places'	käkaṟ 'tiny pieces of stone', käkuṟe 'land' or 'soil'
kan-kal 'waste land'	kañkar 'rubble' used for barren places.
za-dim 'jeweler' , za 'precious stones'	jaṟoä 'jewelry with precious stones', jaruä (Middle Bengali) 'jeweler'
bar-rim-(ma) 'dry land'	bärind 'high and dry land'
ba-ra 'negative verbal prefix'	bärän 'prohibitive'

There are some peculiarities of Bengali that are worth mentioning for their similarities with Sumerian. It is said that in Sumerian a stem is repeated to denote plurarity or intensification of the sense. This is also an important feature of some NIA East Indian dialects, particularly Bengali. For example, *bar̲a* means 'big' but *bar̲a bar̲a gäç* means 'big trees'. An example of intensification of the sense would be *job job* 'very wet / watery' that recalls the Sumerian *zubi* 'water-logged ground'.[8a] Use of reduplicated roots is common, and there are numerous idiomatic expressions, where the words usually function as adverbs or adjectives. Linguists and grammarians assumed such reduplicated words to be onomatopoetic jingles. But originally they could not have been onomatopoetic, because they are always used to convey a specific sense, which led to the discovery that such onomatopes in reality are preservations of older elements. This peculiar characteristic of Bengali onomatopoetics at first sight seems meaningless, but careful comparison brings out that these onomatopoetics are elements from long-gone languages that are also synonymous with the noun or verb which they qualify as an adjective or adverb. This can be illustrated by an example such as *t̲hak-t̲hak käñpä* meaning '(somebody is) trembling *t̲hak t̲hak'*. Here the regular verb to tremble, *käñpä*, is a derivation from Sanskrit *kampan*, while the apparently meaningless pair of words is assumed to be merely onomatopoetic. However, the colloquialism *bhoe t̲hak-t̲hak,* 'trembling in fear', proves that *t̲hak-t̲hak* also has the sense 'tremble', and, in the case of *t̲hak-t̲hak käñpä*, instead of reduplicating the same word, two synonymous words (from two different languages) have been used to intensify the sense. Such words appear to be meaningless, because, in spite of their continuation in idiomatic usage, they have no known etymology. The close resemblance of *t̲hak-t̲hak* with Sumerian *tuku-tuku* 'tremble', and many such examples, prove beyond doubt that the so called onomatopes are legacies of the past. They are so strongly embedded in the language, and provide such a clear image of the action, that Bengali, even after so many centuries of strong influence from Sanskrit, could not shed them. The following examples are very likely preservations of Sumerian words in the disguise of onomatopes or idioms.

Sumerian	**Desi Bengali**
gal 'large, big , great'	gal gal 'excessive, too much'
dara 'continuous flow'	dar dar 'continuous flow of tears, blood'
tuku-tuku 'tremble'	t̲hak-t̲hak (accompanies verb to tremble)
dag / dadag, dag-dag-ga 'to be washed, to be clean'	dag-dage / dag-dag 'to be perfectly clean,'
tál, tál-tál-la 'to make wide, wide spreading'	d̲hal-d̲hal 'too wide' (i.e., 'loose fitting', of clothes)
sir-sir 'serpents'	sar-sar 'describing movement of reptilian creatures'
gaz 'to cut off'	ghac-ghac (accompanies verb to cut)
ti-en, ti-en ti-en 'to be cool, cold'	t̲han, t̲hane 'cool' (in place names); t̲han-d̲a 'cold'
mur 'to crush, to grind'.	mar̲-mar̲, mur̲-mur̲ 'sound of crushing, grinding'
zal-zale 'bright, brilliant'	jal-jale to be bright, shining or glittering,'
za-pa-ág 'noise, sound'	jhapäg 'noise, sound of something falling'

The name of a local god *dängar säñi* (cf. note. 6) worshipped by the boatmen, in Middle Bengali literature is very interesting, because the first word seemingly corresponds to Sumerian *dingir* 'god'. Many of these peculiar characteristics of Bengali and other vernacular languages remain overlooked or explained inadeqately mainly because there are no Near Eastern studies in the Indian Universities, because knowledge about the ancient Near East, its history, culture, and languages is very limited. There is evidence in the epic-puranic texts of close contact between India and the Near East, possibly during the post-Harappan dark period, as this author has found existence of some ancient Near Eastern ethnic names;[9] preservation of some ancient Near Eastern country names in the form of mythical lands;[10] reflection of the kingdom of Mitanni in the myths of the island of the whites;[11] and of ancient Jericho and the river Jordan in the myths of the kingdom Ilavrita and the river Harduni.[12] Recently Semitic words have been found in the Rigveda by C. H. Gordon (note. 7), and Akkadian in Sanskrit by Shendje (note. 3), but Sumerian words although never looked for, are also present in Sanskrit. A few examples here would be relevant: Sumerian *tag* 'to leave, to abandon, to divorce'; corresponds to Sanskrit *tyāga-* with identical sense; and Sumerian *tab* 'to burn, to heat up' corresponds to Sanskrit *tapas* 'heat'. Sometimes Sumerian elements could be hidden, as in *dampati* 'wife and husband, married couple'; here *pati* 'husband' is a known Sanskrit word, but *dam* has no known usage for 'wife' or 'woman'. Moreover, the awkward looking *dam* as a single word unknown in Sanskrit, corresponds to Sumerian *dam* 'wife' or 'husband'. In conclusion, this enquiry is about the need to reexamine vernaculars and ancient Indian texts anew, in spite of many difficulties, in order to determine the exact nature and extent of the contact between India and the ancient Near East.

Notes

1. L. Srinivasan and C. H. Gordon, (1995) "Canaanite Vocabulary in Bengali and in Some Other IE dialects of India," *Mother Tongue*, Issue 1, December.

2. L. Srinivasan, (1996-1997) "papers written in Bengali", published in *Pro Re Nata*, Nos. 27 to 34. Howrah, Calcutta.

3. M. J. Shendje, (1997) *The Language of Harappans: From Akkadian to Sanskrit,* published by Nehru Centre, London.

4. R. Pischel, ed. with critical notes (1938) *The Desinamamala of Hemachandra*, 2nd edition, with notes and glossary by P.V. Ramanujswami. Bhandarkar Oriental Research Institute, Poona, India.

5. S. K. Chatterjee, (1970) *The Origin and Development of the Bengali Language*, George Allen and Unwin Ltd. London. 3 Vols.

6. S. Sen, (1971) *An Etymological Dictionary of Bengali : c. 1000-1800 A.D.*, Eastern Publishers. Calcutta.

7. C. H. Gordon, (1998) "The Near-East Background of the Rigveda," in *Ancient Egyptian and Mediterranean Studies*, in memory of William A. Ward. ed. by L. H. Lesko, Brown University, Providence, Rhode Island. pp.117-120.

8. A. R. Bomhard, (1997) "On the Origin of Sumerian" *Mother Tongue*, Issue III, December. pp. 75-80.

8a. [And cf. Nihali *joppo* 'water', etc. MT-II, and MT-III, p.49. Ed.]

9. L. Srinivasan, (1996) "Mythological Geography of the Puranas". *The Geographical Review of India*, Calcutta. Vol. 58. No. I. pp. 16-26.

10. L. Srinivasan, (1996) " Traces of Near Eastern Names of Land and People in Some Puranic and Mythological Names". *The Journal of the Asiatic Society*, Calcutta. Vol. 38. No. 4. pp. 95-120

11. L. Srinivasan, (1998) "The Mythical Lands of Svetadvipa and Svetavarsa." *The Geographical Review of India*, Calcutta. Vol. 60. No. 1. pp. 42-61.

12. L. Srinivasan, (1999) "The Ilavrita Varsa, Mythical Jambunadi, and Harduni of the Puranas," to be published in *The Geographical Review of India, Calcutta.*

Sumerian and Bengali?
Some Methodological Problems in Long Range Comparisons.

Michael Witzel

Lini Srinivasan's paper exemplarily highlights some of the problems involved with long range comparison. As some of its detractors have maintained for long, one can always find some 50 look-alikes in any two languages. The present paper underlines the difficulties one runs into when one juxtaposes any two vocabulary lists.

While, in a first trial, mistaken look-alikes cannot easily be excluded in the case of modern languages without a long history of written texts, many of them can and should be avoided if the languages involved are attested early. Such is the case with Sumerian (3rd and early 2nd mill. BCE) and the somewhat later (Vedic) Sanskrit (c. 1500 BCE onwards, until today) that precedes modern Bengali by some 3500 years. As the examples given below will indicate, a lot of things have happened between 1500 BCE and 2000 CE.

All of this certainly is old news. Yet I underline these facts, because many of the straightforward comparisons in Srinivasan's paper are easily falsifiable, and such procedure may lend support to those who deny any value to long range comparison. The lists given below underline the value not only of historically attested but also of (internally) reconstructed forms that help to decide on long range comparisons of the word in question. In sum, if we *lack* early attested forms, such as for Australian, we should either proceed with an internal reconstruction of one language or, better, of all Australian languages before long range comparison. Or we should at least proceed, on the one hand, with mass comparison of *all* Australian languages with, on the other, the target of comparison, say Andamanese or Ainu.

In the present context, a mass comparison of modern South Asian Indo-Aryan (IA) languages, from Kashmiri and Nepali to Singhalese, and from Sindhi to Assamese, would have thrown *some* light on the problem. But as we shall see, due to the fairly similar phonological developments in all New Indo-Aryan (NIA) languages, comparisons with the older forms as found in Middle Indo-Aryan (MIA), or Old IA (Skt., Vedic) carry the day.

In addition, one should also pay close attention to words that are not IA or Indo-Iranian (IIr) in shape, such as the fairly stable cluster *-ṇḍ-*, which is 'foreign' to Indo-Iranian (IIr) and Indo-European (IE) but common to Dravidian, among others. This procedure also allows us to identify non-IA sources even if the language(s) of origin are unknown (as they often are; see MT Special Issue 1, Oct. 1999). The example is also useful in so far as it highlights the fact that Sumerian (Sum.) lacks such consonants, a fact which makes direct comparison somewhat hazardous.

The languages neighboring Bengali (and other IA languages) must also be taken into consideration, viz. Tibeto-Burmese in the North, Dravidian in the South and Munda/Austroasiatic west and east of Bengali. They may well have been the source of many of the suspected Sumerian cognates, along with some of the unknown languages just mentioned. In short, the situation is complex and best carried out by team work. However, since most of the words listed below are of IA origin or have been used since Vedic times, an "IA" and "IE" answer from my pen may be useful. The number of Beng. words that remain unexplained, briefly listed at the end of my discussion below, probably can still be lowered (considerably). -- It is a matter of patience and, sometimes, luck to find the proper etymon.

However, as of now, a certain number of interesting words remains, and they could again be tested. They often belong to the indigenous, "*deśī*" category (see below). However, due to the long trade relationship between Mesopotamia and the Indus valley during the Indus Civilization (2600-1900 BCE) and later on --as Greek reports from the time of the Persian empire indicate-- some linguistic connection or loan-word relationship between India and Sumeria should be expected (though I would rather look for it in Sindhi and Gujarati, at the western end of the subcontinent, than at its eastern end, where fewer Sum. words are expected to have survived). In addition, whether the members of the Indus Civilization spoke Proto-Drav. or not, a number of *cultural* loans from Sum. into Drav. can easily be established, (see Blažek and Boisson 1992, and MT, SI 1999). All types of words mentioned so far must be subtracted from a list of inherited vocabulary that is due to a hypothetical genetic relationship between Sum. and the Deśī words of NIA and Bengali. Any Sum. look-alikes in the early attested Vedic (1500 BCE), spoken only in the Greater Panjab, must have entered via another language, for example that of the Indus civilization (MT, SI 1999) and must have been altered by the phonetical peculiarities of that language and of Vedic.

An important question that still has to be answered is why so many words suspected to be of Sumerian origin or to be genetically linked to Sumerian show up only in NIA languages and not before. For, not all of them belong to the sociologically lower levels. During Ṛgvedic and Middle Vedic times, we can expect only little contact with the coast and with Mesopotamia, but from then onwards speakers of IA languages were in maritime contact with Mesopotamia again; if, on the other hand, the so-called Deśī words in NIA go back to pre-IA times (i.e. are pre-Ṛgvedic) the same question as put just now rises again: why does Vedic and even Epic/Class. Skt. show so few traces? (Cf. the appended list of still unexplained words, below).

Another, for the time being rather intractable problem, is the so-called onomatopoetic words which have a structure, a phonetic shape, and frequently a historical transmission of their own that do not agree with the general phonological rules and historical development of the language in question (see below).

Therefore, in order to evaluate the proposals made by L. Srinivasan properly, they have to be viewed in the context of the history of Indo-Aryan, as Bengali is just one of the dozens of modern IA languages. We have a host of sisters of Bengali (Hindi, Nepali, Marathi, etc.) and many predecessors, from Middle and Old Bengali (since the end of the first mill. BCE) to Middle Indo-Aryan (starting c. 500 BCE), Old Indo-Aryan (Classical and Epic Sanskrit, and the earlier Vedic, c. 1500-500 BCE), and Mitanni IA (c. 1400 BCE, in N. Iraq and Syria), and finally their reconstructed ancestors, Indo-Iranian (minimally 2000 BCE?) and Indo-European. To compare Bengali directly with Sumerian brings about all the inherent and well-known problems met with in reconstructing PIE from a comparison of, say, modern Gaelic with Albanian or Armenian.

Since Old Indo-Aryan (OIA) is well attested both in its Vedic and its Epic and Classical Sanskrit form, the situation of evaluating Bengali is similar to that of studying French in comparison with Latin. And the phonetical developments actually are similar as well, e.g. loss of many medial and final consonants or assimilation and simplification of consonant clusters, all of which have the effect of 'telescoping' long inflected words into short, uninflected ones (e.g., OIA *tāvatka* > Beng. *tak* 'up to', or OIA *godhūma* > Beng. *gam* 'wheat'). What might look similar to Sum. *utu* 'sun' in mod. Bengali (*itu, etu*) is not at all so in its predecessor languages (Ved., Class. Skt. *āditya*; see below). Consequently, if the word in question is attested in Class. Sanskrit or in its earlier, Vedic, forms, and if these are dissimilar to Sumerian, this particular comparison has to be discarded. Unfortunately there are, just as in the Canaanite paper in MT I, (see discussion in MT, SI 1999) quite a few such cases.

It certainly is a rather tedious job to go through the dictionaries of NIA (Turner's CDIAL) and OIA (Mayrhofer's EWA, etc.), but it needs to be done before one attempts any comparison between Beng. and Sum.

Many of the still remaining words belong to a category that L. Srinivasan calls *desi.*

Deśī words

Traditional Indian grammar of the middle ages called the MIA words, derived by *historical* development -- as we would say, while they do not! -- directly from Skt. *Tadbhava* ('having their origin from there, [Skt.]') and distinguished them from the more recent, medieval loans from Skt., *Tatsama* ("same as [Skt.]", see Masica 1991: 64 sqq). The rest of the vocabulary, that is *all* that the grammarians could not figure out even with their excellent knowledge of Skt. and Pkt., was called *deśī*, 'local'. As expected, these words include many substrate and adstrate words from the various non-IA languages of the subcontinent. But some of them may also be new formations (cf. Engl. teen speak!) that have no obvious origin in Skt. or Pkt. The grammarians' aim, of course, was (as indeed is Mayrhofer's) to derive virtually everything from Skt. This is allowable in so far as Skt. provides copious materials, but we should not underestimate the range of the linguistic levels *not* covered by texts: words that are likely to come up only in popular speech, such as the famous case of the proper IE *párdati* 'farts', attested only fairly late (EWA II 306), or the vocabulary of farmers, artisans, hunters, etc., which includes also many words that do not come from Skt. or OIA (see below, on *argaḍa*)

It must be emphasized, however, that all Bengali words, whatever their origin, must have undergone the same historical developments as all *explainable* Vedic, OIA or MIA words, as will be outlined immediately below. In fact, any early loan into IA (including all 'subterranean' pre-IA Deśī words that have survived unknown and unseen until modern Bengali) must have undergone the *same* kind of changes. Deśī words, as part of the spoken language, cannot have been exempt from the inherent developments of OIA, MIA and NIA, down to Bengali.[105]

For example, Sum. *da-gal* 'wide, broad' :: Beng. *dāgar* 'large, big, broad', must go back to forms like MIA **daggara* < OIA **dargara*, or Sum. *tur* 'to be small in size, weight' :: Beng. *thorā* 'little, small' < MIA **thavara* < OIA **(s)thavara.*[106] Though in both cases the ultimate origin remains unclear, the reconstructed MIA and OIA forms are unlike those of Sum., and the comparison must be abandoned. If such words were early loans into OIA/Ved., the nature of adaptation of 'foreign' words must be studied (see Kuiper 1991, MT, SI 1999; the same would apply, in their early stages, to words that would have a hypothetical genetic relationship with Sum. and that might have been subsequently changed by MIA and NIA developments).

---

[105] This is overlooked by Srinivasan when she compares the syllable structures of Sumerian and Deśī words.

[106] We find, however, with the opposite meaning, Ved. *sthā́vara* 'fixed' CDIAL 13767, or OIA Ved. *sthaura, sthaula* 'big, large, strong' > MIA: Pāli *thora* 'large' > NIA: Guj., Mar. *thor*, etc. CDIAL 13870; all come from OIA *sthūra, sthūla* 'big', cf. Ved. *sthavīra* 'broad, thick, strong'. However, in this case one may nevertheless assume a special semantic development based on meanings such as Nep. *thoro* 'full grown buffalo heifer', Kumauni *thoro* 'young buffalo bull', *thorī* 'yearling buffalo heifer' > 'young, small'.

MIA and NIA developments

As was seen in the previous examples, the phonetic developments preceding modern Bengali can be summarized as follows. There is a MIA tendency to simplify clusters by assimilation, followed by a NIA tendency to change double consonants to single consonant, often with lengthening of the preceding vowel, and in part, by nasalizing it (see Masica 1991) -- all of which is similar to the well-known developments from Latin > Vulgar Latin > Italian / French: *nox* 'night', acc. *noctem* > *noctẽ* > *notte* / *nuit* [nüi]. There is general loss of final consonants or whole syllables, and there is a 'weakening' of medial consonants, e.g., *t* > *d* > *δ* > *y* > *0*. All of this has a telescoping effect., e.g., OIA *rājaputra* 'king's son' > NIA *raut*, or in place names : *Nāgapura* 'snake town' > *Nagor*, while modern *Nāgpur* is a medieval (Tatsama) loan from Skt. A schematic example of the changes involved is given below.

1500-500 BCE Vedic	500 BCE- MIA/Prākṛt	c.750 AD- early NIA	modern NIA (Bengali, etc.)
hṛdaya 'heart'	hadaya[107]		
	> *haδaya		
	> hiaa[108]	hiyā	(hiyā), (dil)[109]
akṣi 'eye'	akkhi	aṃkh	ãkh
*dargara 'big'	*daggara		dāgar

Procedure

Below, I will follow the order in which the various words are mentioned in L. Srinivasan's paper. I will provide the OIA forms (Vedic, Epic, Class. Skt., or OIA as reconstructed by Turner) and let them be followed, as to show the internal Indian phonological development, by MIA (Pāli, and the slightly later Pkt. forms), and by the NIA forms from N. India (and Sri Lanka): especially, from the more conservative Nepali, and from the central, often more innovative Hindi. For simplicity, Vedic pitch accent (*devá* 'god') has been neglected here as it plays no role in loan words (Kuiper 1991).

Deśī words and Sumerian

- Sum. *giš* 'tree' :: Beng. *gāch* 'tree'.

The word also found in many other NIA languages from Shina (N.Pakistan) and Nepali to Singhalese. It is not attested in Vedic, but found since early post-Vedic/Epic (AV-Pariśiṣṭa) and in MIA: in Pāli (c. 250 BCE) as *gaccha* 'bush'. Turner, CDIAL 3949, reconstructs OIA

---

[107] Found in Pāli.

[108] In the later MIA language Māhārāṣṭrī.

[109] A surface comparison of *hiyā* (older Hindi) and *dil* (loan from New Persian) seems to lead nowhere, but in fact both go back to Indo-Iranian **j'hṛdaya*! They have no relationship with Engl. *heart*, which is found in Vedic Sanskrit as *śrad* + *dhā* = Latin *crē-dō*, 'to put one's heart (Lat. *cor/cordis*), to trust, believe'.

*_gakṣa_, because of the typical distribution of _-ch-_ in the northwest (while the rest should have *_gakkha_). Mayrhofer, _EWA_ III 148 compares Ved. _kakṣa_ 'bush' and deliberates on *_-gakṣa_ as an (unusual!) phonetic variant (Sandhi) in nominal compounds; this would be an IE word (cf. Germ. _Hag, Hecke_). Obviously, Sum. _giš_ and OIA *_gakṣa_ do not make for a good comparison. Theoretically, an adstrate entrance from Sum. is possible as a loan, perhaps via the well attested early oceanic trade; but, _-i-_ > _-a_, and _ś_ > _kṣ_ remain unexplained. There are several early Vedic loan-words that have the, for S. Asia, otherwise unusual cluster _kṣ_, e.g., in the local tree name _akṣoṭa_ 'walnut', a typical tree of the northwestern hills only (cf. MT, SI 1999). In sum: local loan into Skt. (_deśī_ word) from an unknown (NW) substrate language. --- In the sequel, I abbreviate and standardize this extensive treatment, and merely give the derivation from Vedic Skt. > MIA (Pāli, Prākṛt) > NIA (esp. Nepali, Hindi, etc.)

- Sum. giri 'foot' :: Beng. goṛ 'foot', goṛā 'foot of a tree'.

OIA *_goḍḍā_ CDIAL 4272 > MIA: Pkt. _goḍḍa, goḍa_ 'foot'; widely attested in NIA. Not Drav., Munda; no clear source. Retroflexes make it local S. Asian, not Sum. --- Local South Asian _(-ḍḍ-)_ word.

- Sum. gi-gir 'wagon' :: Beng. gāṛi.

OIA *_gāḍḍa_ 'cart' CDIAL 4116, where derived from Vedic _gárta_ (problematic) > Pkt. _gaḍḍa_ etc. 'cart', > NIA Nep. _goṛo_, Hindi _goṛ, goḍā._ Probably, an old word of culture, note Sum. _gígir, gilgul_, Semitic *_gal-gal_, IE *_kek_^{w}_lo-_ > Ved. _cakra_ 'wheel, chariot', Witzel 1999; CDIAL compares Latin _cathedra._ Entry into Vedic(?) or MIA via a local S. Asian form. --- Cultural loan word of Mesopotamian(?) origin, but in S. Asian _(-ḍḍ-)_ form.

- Sum. ge-en, gim 'like, as' :: Beng. jena 'like, as'.

From Ved. _ya-_, pl. _ye_, instr. sg. _yena_ 'by which' > MIA Pāli _yena_, Pkt. _jena_, cf. O. Gujarati _jiṇi_ 'if', thus, 'by which' > 'like'; widely found in NIA, such as in Beng. _je_ 'who, which' (rel. pronoun), see CDIAL 10391. --- Transparent OIA formation < IE pronoun *_yo-_.

- Sum. ga-ar 'wall' :: Beng. ghāro 'wall, room, dwelling' etc.

Skt. _ghara_, only in Buddh. Skt.; Pkt., Pāli _ghara_, > NIA Nep., Hindi _ghar_, from IE *_g_^{w}_horo-_, unlikely, CDIAL 4428; perhaps popular, crossed form of _gṛha_ 'house' (Pāli _gaha_, Avest. _gərəδa_ < IE *_gṛdho_, Gothic _gards_, etc.) and OIA *_gharmya_/Ved. _harmya_ 'house' (Avest. _zairimiia-_ < IIr * _j'harmiya_ < IE?), EWA II 807; cf. also Ved. _āgāra_ 'room, dwelling', late Ved., Pāli, Pkt. _agāra_ 'dwelling', EWA I 159; all(?) these perhaps also crossed with Ved. _geha_ 'house' < *_gaidha_, cf. Avest. _gaeϑa_ 'house, farm', see EWA I 496. --- Probably local IA development of IIr (IE?) form.

- Sum. udu 'sheep' :: Beng. udo 'sheep, stupid'.

Class. Skt. _huḍa_ 'ram' (only in Varāhamihira, c.550 CE, and lex.), Pkt. _huḍa_ 'ram' > NIA only in NW: W. Pahari _huṛ_, hūṛ, W. Panjabi (Lahnda) _huṛeāl_ 'wild hill sheep'; Panj. _huṛiār_ (masc.), CDIAL 14135. KEWA III 541 compares Class. _huḍu, huṇḍa_, lex. _hulu_, "unclear", however, the development _ḍ_ > _ḷ_ > _l_ is normal in medieval Skt. and MIA (see next word), and points to older _huḍu, huḍa._ -- Cf. also the related Ved. _eḍa_ 'sheep' > Pāli _eḷa-ka_, Pkt. _ela_, which survives in NIA only on the rims: in the northwest, in _Sindhi eli-_, and Singhalese _eḷuvā_ 'goat'; note also Class. Skt. _heḍāvuka_ 'horse merchant'; probably from Drav. CDIAL 2512; however, EWA I 264 compares Greek _aiks_, Armen. _ayc_ 'goat'. --- The variations Ved. _eḍa_, OIA _huḍu_, point to a local S. Asian loan-word _(-ḍ-)_ in OIA/Vedic.

- Sum. úr 'thigh' :: Beng. ūru 'thigh'.

Vedic *ūru* > Pāli, Pkt. *ūru* > NIA: Singh. *uruva*, etc. CDIAL 2420; cf. 2421 Ved. *urvastha* 'thigh bone' > Pāli *uraṭṭhi* > NIA, Beng. *urut, urat* 'thigh'. --- EWA I 241 sq.: cf. Iran. **urāva* > Khot. *hurā?* - IE etym. unclear. IIr(?) word; old (RV).

- Sum. uru 'city, town' :: Beng. ūr, Pkt. ūro.

There are two possibilities: First, Epic (and Class.) Skt. *pura* 'fortress, town' > Pāli, Pkt. *pura* > NIA *pur*: Beng. *pur* 'city of Patna'. Note that *-p-* disappears via *-v-* in compound nouns: *Nāgapura* > *Nagor,* etc., or Skt. *rājapura* 'King's town, palace' (Epic) > modern *Rajor* 'Kingston', Hindi *rāwar* 'king's palace', Oriya *rāurār* 'family title' CDIAL 10683; note the same development in 10682 *rājaputra* > mod. NIA *rāut, rāuta* 'a caste, clan', etc.; Ved. *rājaputra* > Pāli *rājaputta*, Pkt. *rāyaütta* [rāya-utta], Beng. *ra(h)ut* 'cavalry soldier; cf. also Skt. lex. *upapura* 'suburb', Class. Skt. *upapurī,* > NIA, only northwest: *wurū́* 'name of a hamlet', CDIAL 2195. --- Better, however, to assume Drav. origin of the Pkt. word, *ūra-* (not *ūro*), which is called a Deśī word in the medieval discussions such as in the collection of Deśī words, the Desinamālā, and given the meanings "village, accociation, multitude'. Drav. *ūr*, Tamil, Kannada, Telugu, Tulu *ūru* 'village, town', Brahui *urā* 'house', etc. DEDR 752. --- Drav. loan, from Sum.?

- Sum. utu 'sun, sun god' :: Beng. itu 'sun god', etu 'sun', Hindi etowar 'Sunday'.

The last item immediately gives the word away: 'Hindi *etowar'*, Nep. *āitavār*, are from *āditya-vāra* 'sun-day', a new formation after the introduction of the Graeco-Babylonian weekdays early in the 1st mill. CE. The word is a rather secondary formation: Ved. *āditya* 'descendent of Aditi; in pl. 'a class of 7-12 gods'; sg. = the sun god', with the typical lengthening in the first syllable (*Vṛddhi*) and *-a* stem suffix ("Vṛddhi derivative") > Pāli *ādicca,* Pkt. *āicca* 'sun', Middle Beng. *āicca* 'a surname'; CDIAL 1153. Skt. *ādityavāra* 'Sunday', found in many medieval texts, CDIAL 1154 Pkt. *āiṃcavāra*, Sindhi *ācaru,* but Nep. *āitavār.* ---Transparent OIA formation *āditya* 'sun' + *vāra* 'day'; *aditi,* probably < *a-diti* 'non-binding', < IE 'non' + dh_1*-ti* 'binding', EWA II 716.

- Sum. uzu 'piest' :: Beng. ojhā 'diviner'.

From Skt. (Manu+) *upādhyāya* (*upa-adhy-āya* 'secondary teacher', from *adhy-i* 'to learn') = 'preceptor, Veda teacher', > Pāli *upajjhāya*, Pkt. *uajjhāa, ujjha, ojjha* > NIA, widely found (wizard, caste name, etc.), Beng. *ojhā* 'snake charmer, exorcist'; CDIAL 2301. --- Transparent OIA formation, *upa* + *adhi* + *ay/i* 'to go close upon (a text), to study', < IE **i* 'to go'

- Sum. mi 'woman' :: Beng. meye, māiā 'woman', M.Beng. māi 'elderly woman, mother'.

Ved. Skt. *mātṛ* (*mātā*, nom.sg.), with regular development > Pāli *mātā,* Pkt. *māyā, māi,* Old Beng. *māva*, Beng. *mā;* forms such as *māi, maiyā* are widespread in NIA; CDIAL 10016. At best, one could compare IE *meh_2ter- with Sum. *m-i*, but baby words beginning with *m-* are, of course, as widely spread in Eurasia as those in *p-* for father (next to *tata, ata* etc.) --- Transparent IE formation, < *meh_2teh_2.

- Sum. mà-a 'I' :: Beng. mui, mai 'I'.

Ved. *aham* 'I'; however, the oblique form *asm-* of the plural *vayam* 'we' is normally used in class. Skt. and NIA for 'I' (i.e. Hindi *ham*, instead of *maī̃*); Skt. *asm-* > Pāli *amh-*, early NIA *amh-* > : Gypsy *ame*, Panj., *asī̃, asā̃,* O. Beng. *ambhe, āmhe, āmi, āmā*, etc.; or rather, from acc. of 'I' > Early NIA *maī̃* > Nep. *ma*, Hindi *maiṃ;* cf. also Ved. *asmākam* 'our' > Gypsy *ama-ro*, Or. *āma-ra*, etc., CDIAL 986. -- Obvious IE. origin: Ved. *aham* < IE **eg'h-om* 'I'

- Sum. im 'mud, clay' :: M.Beng. āmā, 'things made out of burned or unburned clay'.

Perhaps connnected with Ved. *mṛnmaya, mṛttikā, mṛd, mṛdā* 'clay'; however, derived from < MIA **amma-*, < Ved. *arma, arma-ka* 'residue, ruined place, archaeological mound', which involves heaps made of clay, cf. however, EWA I 120. --- Probably IA, old (RV)

- Sum. i-lu 'song, wail' :: Beng. ulu 'ceremonial sound'.

Onomatopoetic: Ved. *ululi, ululu* 'joyous cry', *ululā kṛ* 'to cry out, bemoan' > Pkt. *ulūlu* 'cry of rejoycing'; cf. Latin *ululāre*, IE reduplicated formation, EWA I 230; cf. further Ved. *ulūka* 'owl' --- Obvious onomatopoetic formation, not necessarily IE.

- Sum. dim 'pillar' :: Beng. thām 'pillar'.

cf. MT, SI 1999: *tham* 'to stop' cf. MIA *ṭhape-, ṭhava-* < *sthāpaya-* 'cause to stand, establish', MIA *thāma* < Ved. *sthāman* 'station', cf. Gujarati *ṭhām* 'place' CDIAL 13756-65; cf. also *sthūna / stūnā, stunā*, OP *stūnā* 'pillar', unless it belongs to Ved. *sthūra* 'tall, thick', Avest. -*stura*, Khot. *stura*, thus EWA II 768. --- Obvious IE formation, from $*stheh_2$- 'to stand'

- Sum. dug 'jug, pot' :: Beng. ḍhāk 'large drum', dugi-dugi 'small drum'.

Class. Skt. *ḍhakkā* 'big drum', EWA III 232, onomatopoetic; cf. also similar words, CDIAL 5576 **ḍaḍḍha* 'drum', 5608 *ḍhola* 'large drum', and below, *ṭhak ṭhak*, --- Onomatopoetic formations without clear origins.

- Sum. dur 'bond' :: Beng. dor 'bond, bindings'.

Ved. *dāman*, OIA **dāmana*, 'binding, rope', > MIA **dāmana/dāmara*: Or.: *daũra, daüra* 'rope' CDIAL 6283; cf. also the following. --- MIA, new formation, based on OIA *dā* 'to bind', IE deh_2.

- Sum. dá-ra 'a band' :: Beng. ḍaṛā , ḍaṛi 'a band, rope'.

Ved. *dāman* 'rope' RV+, OIA **dāmana* > Pāli, Pkt. *dāma* > NIA: Hindi *dām*, Beng. *dām(ā), dāmni*; from OIA *dāyayati*, CDIAL 6289 'cause to be tied'. All to be derived from the root Ved. *dā*, pres. *dyati* 'binds' --- IA derivations from IE deh_2 'to bind'.

- Sum. bar 'outside' :: Beng. bār 'outside'.

OIA **bāhira* 'external' > Pāli *bāhira*, Pkt. *bāhira* >NIA, Beng. *bāhir, bār* CDIAL 9226; cf. OIA **bahira* 'external' > Pkt. *bahira* > NIA, Guj. *bahir* CDIAL 9183; all derived from Ved. *bahir* > Pāli *bahi*, Pkt. *bahi, bahiyā* > NIA, M.Beng. *bahi*, Beng. *baï*, CDIAL 9186, EWA II 220; --- Transparent IE formation, < **bheg'h*, > O. Slav., Latvian *bez* 'without'.

- Sum. bara, bara 'seat of a god' :: M.Beng. bārā 'enclosed spot of a deity'.

OIA **vārtra* 'pertaining to a fence' > Skt. (Epic, re-borrowed < MIA) *vāṭa* 'enclosure, fence', Class. Skt. *vāṭikā* 'garden'; Pāli *vāta* 'enclosure, circle', Pkt. *vāḍa* 'fence', > NIA, widespread: Beng. *bāṛa* 'edge, border, selvedge of cloth', Nep. *bār* 'hedge, boundary of field', etc., CDIAL 11565, CDIAL 11480. All derived from Ved. *vṛ* 'to enclose'; note Ved. *vartra* 'dike, dam', but Nur. *watrwo, wəṭë* 'cattle shed', CDIAL 11370, cf. also Ved. *vārāyate* 'obstructs, keeps back', Pāli *vāreti*, Pkt. *vārei*, widely found in NIA; note Beng. *bārā* 'to shield, stop', Hindi *barnā* 'to forbid', Nep. *barnu* 'to limit, surround, give up, refrain from', etc., CDIAL 11554, EWA II 512. --- Transparent formation, from Ved. *vṛ*, pres. *vṛṇoti* 'enclose', Iran. **var*, < IE **wel.*

- Sum. pal 'turn of office' :: Beng. pālā 'turn of office'.

OIA *paryaya* 'revolution, turn' > Pkt. *pajjaya, pallava* 'succession, order', Beng. Hindi *pālā* 'turn', CDIAL 7937. --- Transparent OIA formation: *pary* 'around' + *i* 'to go', Ved. dial. *palyi* 'to go around' < IE **peri + i.*

- Sum. sar 'whole, all' :: Beng. sārā 'whole, all'.

Obviously connected with CDIAL 13276, Ved. *sarva* 'all' > Pāli *sabba*, Pkt. *savva* > early NIA *sava, sāva, sahu, sāhu* > NIA: Nep., Beng. *sab*; NIA: note the extension *-ḍa*: Gypsy *savoro, saró* etc. --- Transparent IE form, **sol(H)wo, Greek hólos*, Latin *saluos, saluus.*

- Sum. śá-an 'clever' :: Beng. seyanā, seanā, seyān 'clever'.

Unlikely, connection with Ved. _sayana_ 'binding', _śayana_ 'lying, sleeping' --- Unclear.

- Sum. rin-na 'oven' :: Beng. rānnā 'cooking'.

Ved. and Class. Skt. (comm.) _randh_ 'subdue, torment'; 'to cook' in Late Ved., Class. Skt.; for details of the development of meaning, see CDIAL 10616; EWA II 431, Iran: Khot. _haran_ 'throw'; --- Probably < IE _*lend_ 'to slope down'.

- Sum. nitah 'male' :: Beng. nitā 'male worker.' --- Unclear
- Sum. tum 'abundance, plenty' :: Beng. dhum 'plentiful, abundant.' Probably 'onomatopoetic'. Unclear.
- Sum. lu-gal 'king' :: Beng. (E. Pkt.) laulā 'king', raulā (M.Beng.) 'king, royal'.

See above _rājaputra_; the forms with _l_ are dialect variants already in MIA (eastern _lāje_, western _rājo_); EWA II 444 sqq. --- Obvious IE word, from _*h_$_3$_rēg'-s_ 'ruling, chieftain', Latin _rēx_ 'king', etc.

- Sum. lu 'man, fellow' :: Beng. loa (O.Beng.) 'man, fellow'.

Ved. _loka_ 'free space, clearing, world' > Pāli _loka_ 'world', Pkt. _loga, loa_ 'world, people', Hindi _loe_ 'people', etc.; CDIAL 11119; --- Obvious IE formation, _*leuko-_, Latin _loucos, lūcus_ 'clearing', Dutch (water-)_loo_, Lith. _laũkas_ 'field', etc.

Deśī root Nouns and Verbs

- Sum. qar 'to take away' :: Beng. kāṛ 'to take away by force'.

From _kṛ_ 'to do, make' CDIAL 3057?; cf. Epic Skt. _kara_ 'tax', Class. Skt. (Pāṇ.) _kāra_ 'tax', Beng. _kar_ 'land revenue' CDIAL 2780; probably, a special semantic development, from _kṛ_ 'to do, make' (EWA III 59)?, or rather, from Drav.? Burrow compares Tamil _kaṟa_ 'to milk', _kaṟavu, kaṟai_ 'milking, tribute' DEDR 1385. --- From IE or Drav.

- Sum. gar 'to do, make' :: Beng. gaṛā 'to do, make'.

Ved. _kṛ_ 'to do, make', 3rd sg.. _kṛṇoti, karoti_ > Pāli _karoti_, Pkt. _kareï, karaï_ > NIA: Hindi _karnā,_ Nep. _gar-nu_!, etc. Beng. _karā_. Rather, from OIA _*gaṭhati_, MIA Pkt. _gaṭhati_ 'makes' > Beng. _gaṛā,_ Hindi _gaṛhnā_ 'to hammer into shape, form', O.Awadhi _gaḍhāi_ 'makes', CDIAL 3966. --- Local South Asian _(-ḍh-)_ substrate word.

- Sum. gúr 'to bend down, bow' :: Beng. gaṛ 'to bend down to show respect').

Compare the semantically difficult CDIAL entries 3973 _gaḍati_ 'drops, distils, runs as a liquid', _*gaḍḍ_ 'to dig, bury', 3967 'ditch', 3986 'fort'. However, the reconstruction from OIA/MIA _*gaḍ(h)_ indicates local S. Asian origin. --- Unclear: probably local _(-ṛ-)_ South Asian loan.

- Sum. gur 'to grind, to rub down' :: Beng. gurā, guṛo 'to grind, to powder down'.

Class. Skt. _guṇḍaka_ 'dust powder' lex., Pkt. _guṃdana_ 'smearing with dust', _guṃdia_ 'covered with dust' > Beng. _gūṛā_ 'powder, pounded'; cf. also Pkt. _guṃṭhaï_ 'makes dusty', CDIAL 4193, EWA III 159. --- From Drav., Telugu _guṇḍa_ 'powder', etc. DEDR 1692

- Sum. gúr 'to twist, roll up' :: Beng. gaṛā, gaṛo 'to roll'.

Skt. _guḍa_ 'globe, ball' > Pāli _guḷa_, Pkt. _gulia_ > Beng. _guṛ_ 'globe, mouthful' CDIAL 4182; _guli_ 'ball, pill' EWA III 160, _guḍa_ III 158, _gola_ 'ball, globe', "probably loan word" III 165. --- Local S. Asian loan word _(-ḍ-)_.

- Sum. gu-tab 'to twist yarn' :: Beng. gutā 'to twist yarn, to roll up things'.

Cf. Drav. group of DEDR 1713 sq., Tam. _kuntai_ 'loop, running knot', Kan. _kude_ 'fetter', _gudi_ 'to tie feet or legs', Tel. _kudigu_ 'to become tight (knot)', etc. --- Unclear; Dravidian?

- Sum. bad 'to separate, divide, part' :: Beng. bād 'to separate, discard'.

Probably from *vārayate* 'to obstruct, keep back' CDIAL 11554 (see below, sub: Sum. *ba-ra*). --- Probably of IA, IE origin.

- Sum. par 'to go, pass by' :: Beng. pār 'to pass by, to go past, to go across'.

Ved. *pārayati* 'brings over' > Pāli *pāreti* 'makes go through', Pkt. *pārei* 'crosses, completes' CDIAL 8106; cf. *pātayati* 'makes move quickly, throws' > Nep. *paṛnu*, Beng. *pāṛā* 'to fell, lay down' CDIAL 8053, EWA II 85 sq. --- Transparent IA formation from IE origin.

- Sum. pá(d) 'to call, to speak' :: Beng. pāḍ in dāk pāḍ 'call', hā pāḍ 'to shout' --- Onomatopoetic? unclear.
- Sum. pad 'to tear out, remove' :: Beng. pāḍ 'to tear out' --- Unclear. Local South Asian *(-ḍ-)* substrate word.
- Sum. men 'to agree, allow' :: Beng. mānā 'to agree, concede'.

Ved. *mānayati* 'makes think, esteems, honors' > Pāli *māneti* 'honors' > NIA: Sindhi *māṇaṇu* 'enjoy' etc. 10047; Ved. *māna* 'wish, pride' > Pāli *māna* 'pride', Pkt. *māṇa* 'pride' > NIA: Nep. *mān* 'respect', Beng. *māṇ* 'pride' CDIAL 10040, EWA II 305 *man* 'to think'. --- Obvious OIA formation < IE **men* 'think'.

- Sum. nad/nu(d) 'to lie down, sleep' :: Beng. nid 'to sleep'.

Ved. *nidrā* 'sleep' > Pāli *niddā* 'sleep', Pkt. *ṇiddā* > NIA: Hindi *nīd* 'sleep' etc.; EWA II 757. --- Transparent OIA formation: *ni* 'at the proper place' + *drā* 'to sleep'; < IE **dreH*, cf. **dr-em* (Latin *dormīre*), etc.

- Sum. nir-nir 'winnow' :: Beng. naṛ, naṛ-naṛ 'to move, shake'.

OIA **naṭati* 'trembles, totters' CDIAL 6934, EWA III 282, *naṭati* 'dances' < Ved. *nṛt*; cf. also CDIAL 6933 Epic *naṭa* 'dancer', Pāli *nata*, Pkt. *nada* > NIA: Beng. *naṛ* 'caste of dancers', Beng. *nāṛā* 'to shake, flinch, move, shake, agitate'. All from Ved. *nṛt* 'to dance', Nur. Ashkun *nāṭ;* ultimate etymology 'unclear' EWA II 21. Note that the OIA form (old Vedic Sanskritization of *naṭ*?, cf. in general Kuiper 1991) differs from Sum. --- Old IA formation, IE(?) etymology unclear.

- Sum. šár 'to write' :: Beng. sarā 'to write' in: kalam-sarā 'to write with a pen'.

Ved. *sārayati* 'makes move, remove' CDIAL 13358 (see next); *kalam* 'pen' < Arab. < Greek *kálamos* 'writing reed, pen' EWA III 71. --- Transparent MIA formation < IE root **sal.*

- Sum. sar 'to drive away, forward' :: Beng. sarā 'to move things, to remove, drive away.'

Ved. *sārayati* 'makes move, remove' > Pāli *sāreti,* Pkt. *sārei* > NIA: Nep. *sarnu* 'to move' Beng. *sārā* 'to repair, finish, ruin' CDIAL 13358. --- Obvious OIA formation < IE **sal* 'start running', Latin *salire* 'to jump'.

- Sum. si 'to lay' :: Beng. šo 'to lay down, lie down, sleep'.

Ved. *svapati* 'sleeps', OIA **supati* > Pāli *supati*, Pkt. *su(v)aï* > NIA: Panj. *sauṇā,* Beng. *soya* 'to sleep, lie down' (note, general Beng. pronunciation of *s* = *š*); EWA II 791. --- Obvious IE formation, **swep* 'to sleep', Latin *sōpīre* 'to fall asleep', etc.

- Sum. til, til-til 'to pick off, pluck' :: Beng. tolā 'to pick off, pluck'.

Perhaps connected with OIA *tolayati* 'weighs', Beng. *tolā* 'to weigh', CDIAL 5979. --- Unclear.

- Sum. te(gá) 'to attain, reach, touch upon' :: Beng. thekā 'to reach, to touch upon, to meet'

Ved. *stheya* 'to be stationed' > Pkt. *thea* 'durable' > NIA: Beng. *thekā* 'to be checked', CDIAL 13777, EWA II 764 sqq. --- Obvious OIA formation, from *sthā* 'to stand', < IE $steh_2$ 'to stand', Greek *hísthēmi*, Latin *statio,* German *stehen*, etc.

- Sum. tag 'to spoil' :: Beng. ṭak 'sour, spoiled food.

Ved. *takra* 'buttermilk' > Pāli, Pkt. *takka* 'buttermilk' > NIA: Beng. *tak* 'sour, acid', *ṭakā* 'to grow sour', *ṭak(a), ṭakuyā* 'sour' CDIAL 5617, EWA I 614-5; from Ved. *tañc* 'to contract, coagulate'; cf. Icelandic *ϑēl* 'butter milk'. --- Obvious IE root **tenk* 'to coagulate', Lith. *tánkus* 'dense', etc.

- Sum. rar(r) 'to cut', dar 'to split' :: Beng. thoṛ, thoṛā 'to chop, rend'.

Cf. Skt. *truṭ* 'to break', *truṭyati* 'is broken', 6098 **thuḍa* 'trunk'; see next entry; EWA III 257, s.v. *troṭ* 'to break', etymology "unclear". --- Local S. Asian loan *(-ṭ/ṭh-).*

- Sum. tar(r) 'to break' :: Beng. toṛ, toṛā 'to break'.

Skt. *truṭ* 'to break', *truṭyati* 'is broken', OIA **troṭa* 'breaking, breakage', Pkt. *toḍa* 'breaking' > NIA: Hindi *toṛ* 'breakage' CDIAL 6077, EWA III 257, s.v. *troṭ* 'to break', ultimate etymology "unclear" --- Local S. Asian loan *(-ṭ/ṭh-).*

- Sum. tuk 'to have' :: Beng. thāk 'to have'.

OIA **sthakk* 'stop, halt', Pkt. *thakka* 'stopped, remaining, tired', Assam. *thākiba* 'to stay', Beng. *thākā* 'to be at a standstill, exist, be' CDIAL 13737; --- Obvious IE root, MIA development.

Aspirated/nonaspirated consonants

- Sum. gur 'to run about' :: Beng. ghurā 'to travel, to travel around in a place'.

OIA **ghurati* 'revolves', Pkt. *ghulaï* 'turns' > NIA: Beng. *ghurā* 'to roll around' CDIAL 4488; cf. CDIAL 4497 Epic *ghūrṇati* 'moves to and fro'; < Drav. according to Burrow; cf. CDIAL 4526 *(pra-) gholayati* 'mixes, stirs together'. Note that the (here, hypothetical) interchange of aspirated and non-aspirated velars (*k/kh, g/gh*) in OIA And NIA (Srinivasan: *gur/ghur*) seem to reflect assimilation of foreign words in to Vedic (and later), as explained by Kuiper 1991. --- MIA word of unknown origin.

- Sum. gur-gur 'to reel' :: Beng. ghurāṛ, ghorāna 'to reel' --- See preceding.
- Sum. gur 'to return, come back' :: Beng. ghur (with aux.) 'to return, come back, give back' --- See preceding.
- Sum. guz 'to crush, smite' :: Beng. ghuc, ghucā 'end, destroy with force or violence'.

**ghūssa* 'blow with fist', only NIA: Sindhi *ghūso* 'fist', Nep. *ghussā*, Or. *ghusa* 'blow with fist', Beng. *ghus*, Hindi *ghūsā, ghussā* 'fist', Guj. *ghusɔ*, Marathi *ghusā, ghussā* CDIAL 4498. OIA **-ss-, *-Cs-, *-sC-*. Probably a loan from Munda: Santali *ghusa, ghusa ghusi* 'to prod with the finger passed between the fore and middle finger', *ghus ghus* 'internally'. Or both Munda and IA from a local substrate. --- Late(?) loan word from Munda.

- Sum. gir/ ger 'to enclose, to grid' :: Beng. ghirā/gherā 'to enclose, to grid, to surround'.

OIA **ghir, gher* 'make go around, surround', only found in NIA: Nur. (Waigali) *girum* 'whirlpool', Hindi *ghirnā* 'to be surrounded', Nep. *ghernu*, Hindi *ghernā* 'to surround'; from Drav. CDIAL 4474; cf. Tamil *kuruḷ* 'to curl', DEDR 1794, and Skt. *kurala* 'curl'. --- from Drav.? Unclear, see above Sum. *gur*, NIA *gur/ghur.*

- Sum. ku 'to eat' :: Beng. khā 'to eat'.

Ved. *khādati* 'chew, bites' > Pāli *khādati* 'eats', Pkt. *khāaï, khāi* 'eats' > NIA: Nep. *khānu*, Hindi *khānā* 'to eat' CDIAL 3865; EWA I 451, cf. N.Pers. *xayidān* 'to chew', Baluchi *khāδay* 'to eat'. --- Obvious IIr. root, not from Sumerian, but maybe from a Central Asian language; IE origins unclear.

- Sum. kuš 'skin, hide, leather' :: Beng. khosā 'skin (of vegetables)', khos 'scab, skin disease', khuski 'dandruff'.

Skt. (lex.) *khasa* 'scab' > NIA: Guj. *khas* 'itch, scab', CDIAL 3854, EWA III 143; cf. OIA **skoṣati* 'plucks out, pokes', from Ved. *sku, skunoti* 'pokes', *skuṣṭvā* 'having poked, removed', *ā-sku* 'to cut marks' > NIA: Nep. *khosnu* 'to snatch away', Hindi *khosnā* 'to pluck out', etc.

CDIAL 13661; EWA compares Iran. Khot. *bu-ṣkuta* 'torn' etc., < IE. **skeuH* 'to poke, cut into', Hittite *iskuna(hh)* 'to mark, designate' (by cutting marks), Latin *cūtis* 'skin', German *Haut*, etc. --- Cf. further Class. Skt. *kacchū* 'itch, scab', Pāli, Pkt. *kacchu, kacchū* > NIA: Or. *kācchu*, Singh. *kas*, CDIAL 2621, Class. Skt. *kacchu-ra, kacchu-la* 'affected with itch', Pkt. *kacchuria, kacchula*, Or. *kāchurā*, CDIAL 2620 sq., 2748; EWA III 44sq. compares Drav.: Kannada *kajji, gajji* 'itch'. --- However, note especially the old word, Ved. *kaṣ* 'to cratch' > Pkt. *kasaï* 'rubs' > NIA: Or. *kasibā* 'to rub, scrape, scratch', Hindi *kasnā* 'to assay, test,' Guj. *kasvū* 'to test by rubbing', etc. CDIAL 2972; EWA I 331 sq. with IA etymology, popular form of *karṣ* 'to scratch; (and thus IIr, from IE **k*^{w}*el*, EWA I 319 sq.), which might explain this word, but not the forms *khasa, kacchū*. --- The divergent forms *kaṣ, khas, khacch* rather indicate some old, mutually independent words, some (*kacchū* etc.) from a substrate or adstrate. --- Sumerian is not involved.

- Sum. tibi 'palm, a blow with the palm of the hand' :: Beng. thābrā , thābar, thāppar 'patting, slapping with the palm'.

OIA **ṭhapp, *thabb, thipp* 'slap, pat' > NIA. Kumauni *thāp*, Hindi *thāpnā* 'to slap'; Or. *thābā* 'paw', etc. -- See Kuiper 1948: 60, Santali *ṭhapa ṭhobo* 'to slap, beat slightly, to cuff', Mundari *thābri*; Ho *caprā* 'to slap' > Class. Skt. *capaṭa, capeṭa* 'slap in the face', see CDIAL 4673 etc. Obvious loan from Munda, cf. DEDR 2335 for loans in Drav. (Tam. *cappāṇi*, Telugu *capaṭa*, etc.); cf. above, Sum. *guz* :: Beng. *ghuc, ghucā* for a similarly wide spread word. --- Probable loan from Munda or another S. Asian source.

- Deśī çut (= cut) 'to run' ~ Hebrew šūt 'to go to and fro'.

Ved. *cyavate* 'moves to and fro' (Y.Avestan *šavaiti*, Khot. Saka *tsu*, M.Pers. *šudan* 'to go', EWA I 552-3 < IE **kyew*) > Pāli *cavati* 'moves', Pkt. *cavaï, cayaï* > NIA: O.Gu. *cavaï* 'falls', Singh. *senavā* 'to fall' DIAL 4939. What is Beng. *cu-t*? --- Old IIr and IE verb.

- Deśī çiṛ, çeṛ 'to tear' (= ciṛ), çeṛā 'rags, tatters' ~ Sum. túg-mu-sir-ra 'rags, tatters'.

Ved. *cīra* 'strip (of bark, cloth') > Pāli, Pkt. *cīra* > NIA: Nep. *ciro* 'cut, slice', Hindi *cīr*, Beng. *cir* 'tear, slice, lath, old torn cloth', etc. < Drav. acc. to CDIAL 4843, EWA I 392, cf. DEDR 2491 Mal. *cintuka* 'to tear', Tel. *cīru* 'to gash, rend, tear', Kurukh *cīrnā* 'to divide by rending, etc;' Brahui *cīṛing* 'to slite, slice, split open', etc.; but cf. also Munda: Santali *cira, ciro* 'to tear, rend, split; a strip', *ciri biri, ciri biti* 'in small pieces, strips', *cirkau* 'to crack; EWA I 545 compares Khot. *cile* (pl.) 'dress' < **kilo* (?); cf. also OIA **cīrayati* 'splits' CDIAL 4844. --- Perhaps IIr, otherwise Late Vedic < Drav./Munda?

- Sum. šu 'hand; to do things with the hand' :: Deśī ço, çu (śo, śu) 'to touch, feel with the hands'; and from the same Deśī word: śur 'tentacles, trunk (used like hands)'.

Epic Skt. *śuṇḍa* 'elephant's trunk' > Pkt. *suṃḍā* > NIA: Nep., Hindi *sūṛ*, cf. Or. *suṇḍa* 'trunk, tentacle', CDIAL 12516; EWA III 493 'foreign word'. However, cf. also Ved. *tuṇḍa* 'snout, trunk', CDIAL 5853, Kuiper 1948: 152sqq., KEWA I 510. --- Local South Asian *(-ṇḍ-)* loan.

Sumerian and Deśī Words

- Sum. a-aga 'instruction, degree', a-ag 'to give command to' :: Beng. aggā 'instruction, order, command'.

A secondary loan (Tatsama, with typical substitution of *jñ* by North Indian *gy*), from Skt. *ā-jñā* 'command' *ājñāpayati* 'directs, orders'; direct derivation (Tadbhava) > Pāli *āṇā* 'command', Pkt. *aṇṇā, āṇā* 'order' > NIA: Hindi *ān* 'order', Or. *āṇ* 'resolve, wish, anger'; Pāli *āṇāpeti* 'orders', Pkt. *āṇave, āṇvai* 'orders' > NIA: Or. *anāība* 'to take care of, wait, look', CDIAL 1095-

6; EWA I 599sqq., < IE _*g'neh_$_3$, Greek _gi-gnō-skō,_ Latin _(g)noscō,_ etc. --- Obvious IA formation based on IE.

- Sum. aga 'crown, tiara, i.e. things worn on top of head' :: Beng. āgā 'top or tip of things'; M.Beng. 'head, top'.

Ved. _agra_ 'top, summit' > Pāli, Pkt. _agga_ 'top, front, point' > NIA: Hindi _āgā_ 'front, forehead', Singh. _aga_ 'top', etc. CDIAL 68; EWA I 45, compares Avest. _aγra_ 'first' etc., no clear IE etymology (Latvian _agrs_ 'early', _agrums_ 'early time')? --- Old IIr. word, perhaps even IE. - Not Sumerian.

- Sum. ér 'water' :: Beng. irā (Pkt.) 'water'.

Tatsama loan from Skt.: Ved. _irā_ 'drink, refreshing draught', cf. CDIAL 14296, Epic _Irā-vatī_ 'having refreshing drinks, a Panjab river', Panj. _Rāvī;_ EWA I 195, 187 compares, with complex relationships, Ved. _iḍ, iḍā,_ Avest. _īža,_ Ved. _iṣ_ 'refreshing drink' EWA 198, 271, all ultimately from IE _*[h_$_1$_]eis-[h_$_2$_],_ Greek _īnáo, iaínō,_ Latin _īra,_ O.Norse _eisa,_ etc. --- Obvious IE origin.

- Sum. ig 'door' :: Beng. āgal (poetic) 'door'.

Ved. _argaḍa, argala_ 'bolt' > Pāli _aggaḷa_ 'cross bar at a door, bolt', Pkt. _aggala_ > NIA: Nep. _āglo_ 'bolt, Or. _āgaṛa_ 'gate', Beng. _āgaṛ, āgoṛ_ 'mat door, hurdle', Hindi _āgal,_ etc. CDIAL 629; because of the unusual word formation _arga-ḍa,_ EWA I 114 assumes non-IE origin, and compares Munda words. DEDR, appendix 9 < Skt.! --- Sumerian differs sufficiently from the older forms (arg-).

- Sum. bui 'knowledge, learning' :: Beng. bai 'book (of knowledge)'.

Class. Skt. _vahikā_ 'list, book' > Pkt. _vahiyā_ 'accounts' > NIA: Nep. _bahi,_ Or., Beng. _bahi, baï,_ Hindi _bahī,_ etc. CDIAL 11460; from Skt. _vahati_ 'carries, is carried along' > Pāli _vahati_ 'carries, proceeds', thus: 'proceedings', as in _rāja-vāhikā_ 'kings diary'? --- IA (and IE) origin likely: IA _vah,_ IE _*weg'h_ 'to drive'.

- Sum. gun 'tribute, talent' :: Beng. guna-gār 'to pay/compensate money'.

Probably from Ved. _guṇa_ 'thread, strand of cord (see CDIAL 4190), quality, x-fold, x-times value', as in Class. Skt. _guṇa-kāra_ 'multiplier', _guṇita_ 'multiplied', thus _guṇī_ + _kṛ_ 'to multiply' > 'compensate? Note OIA _guṇāyati_ > Pkt. _guṇei_ 'counts' > Beng. _gunā_ 'to count'. IA _guṇa_ is of doubtful IIr origin, probably local South Asian _(-ṇ-)_ loan; cf. EWA I 498. --- Unclear, perhaps IA.

- Sum. sir 'to belong' :: Beng. cir 'very longtime, life long, forever'.

Ved. _cira_ 'long, lasting' > Pkt. _cira_ 'long' > NIA: Panj. _cir_ 'space of time, delay', Guj., Mar. _cir_ 'for a long time', etc. CDIAL 4824; EWA I 544: etymology not very clear, perhaps _ci-ra_ 'heaping up, piling up', from IE _*k_^{w}_ei_ 'to heap up' EWA I 532, cf. parallel formation in O.Iran.: Avest. _vī-cira_ 'piling separately, distinguishing.' --- Old attestation (RV+, Avesta) points to IIr origin.

- Sum. kur 'mountain' :: Beng. cuṛ, cuṛā 'peak of mountain'.

Ved. _cūḍa_ 'topknot, protuberance' > Pāli _cūḷa_ 'protuberance', knot', Pkt. _cūḍā, cūlā_ > NIA: Nep. _cur_ 'tenon', _culi_ 'mountain peak', Hindi _cūṛ_ 'top knot', etc. CDIAL 4883; EWA I 546 compares Avest. _ašta.kaožda_ 'name of a diadem', but also evaluates non-IIr. origins: CDIAL 4883 _*coṇḍa, *coṭṭa, *cunda_ point to several loans from a substrate, by necessity differing from Sumerian (_k :: c_), such as Drav. _cūṭu_ 'to carry on the head' < Munda : _cūṭa_ 'tuft of hair', etc., cf. Kuiper 1948: 154 --- Local S. Asian _(-ḍ-)_ substrate origin.

- Sum. kar 'town' :: Beng. gaṛ 'fortified town'.

OIA _*gaḍha_ 'fort', Pkt. _gaḍha_ 'fort', Panj. _gaṛh,_ Or. _gaṛ(h)a,_ etc. 'fort'; cf. also _gaḍa_ 'ditch' CDIAL 3967. --- Local South Asian loan _(-ḍh-)._

- Sum. tag 'to weave' :: Beng. ṭāku 'spindle of a loom'.

Ved. *tarku* 'spindle' > Pkt. *takku* > NIA: Or. *ṭaku*, etc. from *tark* 'revolve', CDIAL 5717 --- Transparent IE origin, from **terk*(^{w}) 'to turn', Latin *torquēre*, Hittite *tark*, etc.

- Sum. ku, kua 'fish' :: Beng. keut 'fisherman' (M.Beng.), keorā 'a caste of fishermen'.

Ved. *kevarta, kaivarta* 'fisherman' > Pāli, Pkt. *kevaṭṭa* > NIA: Hindi *kewaṭ* 'fisherman, boatman' etc., CDIAL 3469, EWA I 402 "unclear, loan word?" cf. Witzel, MT SI, Oct. 1999. --- Old local loan word (RV+).

- Sum. kemu 'flour' :: Beng. gam 'wheat'.

Ved. *godhūma* 'wheat' > MIA: Pāli *godhūma*, Pkt. *gohūma* > NIA: Nep. *gahũ*, Hindi *gahũ* etc. CDIAL 4287, EWA 498, see MT, SI 1999: < O.Iran. *gantuma*, Hittite *kant*, etc. --- Popular etymology 'cow smoke', based on Near Eastern word of agriculture (see MT, SI 1999).

- Sum. sur 'to squeeze, to press out ' :: Beng. šuṛi 'the caste name of the wine maker and seller'.

Ved. *śuṇḍa* 'pub', Class. Skt. *śuṇḍin* 'preparer of spirituous liquor' > Pkt. *saumḍia* > NIA: Nep. *sũṛi* 'a caste in the Tarai', Or. *suṇḍi*, etc. CDIAL 12519, EWA I 464, etymology "unclear". --- Local SouthAsian *(-ṇḍ-)* loan.

- Sum. sila 'measure of capacity' :: Beng. sera 'a traditional unit of capacity'.

OIA **satera* 'a measure of weight'; Pkt. (Gāndhārī in Niya, Xinjiang!) *satera* > NIA: Nep. Hindi *ser*, etc. CDIAL 13106, EWA 485 compares *śadaka* 'chaff'(?) --- Unclear.

- Sum. iku unit of area' :: Beng. ekar 'traditional unit of area'.

Probably related to Ved. *eka* 'one', as in *ekatara* 'one of two' > Panj. *ikerā* 'once only', *ekatva* 'oneness, unity' > Singh. *ekatu* 'united', etc., thus **eka-kāra* 'doing it, counting it one time'. --- Unclear.

- Sum. izi 'fire' :: Beng. āc 'fire'.

Ved. *arcis, arci* > MIA: Pāli *acci, accī,* Pkt. *acci* 'heat' > NIA: Nep., Hindi *āc*, etc. CDIAL 635; EWA 114, IE **$(h_1)erk^w$* 'to shine'. --- Transparent IE word.

- Sum. tur 'to be small in size, weight', tur-ra 'small, childish' :: Beng. thorā 'little, small'.

From MIA <OIA **(s)thavara*, cf. Ved. *sthaura*, which would disagree with any Sum. link; see above. --- Unclear.

- Sum. guru 'a large measure of grain' :: Beng. guru 'adjective for heavy weight'.

Tatsama from Ved. *guru* 'heavy', Class. Skt. *guru* 'heavy, Guru' > MIA: Pāli, Pkt. *garu* > NIA: Nep. *garũ*, Or. *garu*, Hindi *garuā* 'heavy, venerable', etc., CDIAL 4209; EWA I 490, IE *g^wṛh_2-u*, Greek *barús*, Latin *gravis*, etc. --- Transparent IE word formation.

- Sum. gur 'heavy, thick' :: Beng. guru 'heavy, serious, important'.

See preceding: Tatsama from Ved. *guru* 'heavy'. --- Transparent IE word formation.

- Sum. ma-lah/ má-lal-a 'boat man, sailor' :: Beng. mālā, mallā 'caste of mariners, sailors, fishermen'.

Cf. Ved. *Malla* 'wrestler, a tribe, mixed caste' (note also *jhalla* 'wrestler'), cf. Beng. *mājhi* 'a caste of boat rowers and boatmen' (below), possibly from **ma-jjh-*, as *ma-* can be a Munda prefix indicating possession; for the latter, cf.(?) also *jhaṣa, caṣa* 'a large fish' (MT SI 1999), --- Unclear.

- Sum. ^{giš}má-šu (-a) 'ship, deep draught boat' :: Beng. mājhi 'a caste of boat rowers and boatmen'.

OIA **majjika* 'boatman' (cf. *maṅga* 'head of boat'?, CDIAL 9705, EWA III 377) > NIA: Nep. *mājhi*, Or. *mājhi*, Hindi *mājhi*, CDIAL 9714; any connection with Ved. *matsya*, Buddh. Skt., Lex. *maccha* 'fish'? Rather, with Ved. *majj* to 'dive under, drown' > NIA: Beng. *majāna* 'to immerse', O.Hindi *majjanaüṃ* 'to sink, bathe', etc. EWA II 291, < PIE **mesg* ~ PUral. **mośke* 'to wash'. --- Unclear, perhaps local word.

- Sum. mun 'salt' :: Beng. nun 'salt'.

Ved. *lavaṇa* 'salt' > Pāli, Pkt. *lavaṇa* > NIA: Panj. *lūṇ, nūṇ* (with assimilation *l...n* > *n...n*), Nep. *nun*, Or. *luṇa, nuṇa*, etc. CDIAL 10978; EWA II 475 from Ved. *lu* 'to cut'? This from IE **leuH* 'to cut off'. At any rate, the Ved. forms do not fit Sum. (*lavaṇa : nun*). --- Perhaps IE; otherwise local South Asian *(-ṇ-)* substrate word .

- Sum. na-ru-a 'stela', Akkad. naru 'stone for monuments, boundary stones' :: Beng. noṛā 'term for round and elongated grinding stone'.

Class. Skt. *loṭha* 'rolling' > Pkt. *loḍha* 'rolling pin' > NIA: Nep. *lohoro* 'stone rolling pin', Beng. *loṛā, noṛā, nuṛi* 'rolling pin' CDIAL 11134, cf. **lortati* 'rolls' CDIAL 11157. Same change of *l/n* as in *lūn*, above. --- Probably local South Asian *(-ṭh-)* loan word.

- Sum. zal 'to flow'; zalag 'to cleanse, purify' :: Beng. jal 'water'.

Ved. *jala* 'water' > Pāli, Pkt. *jala* > NIA: Nep., Hindi *jal*, etc.; note Nahali *jappo* 'water'; CDIAL 5155; EWA I 579 perhaps < IE **g^welH* 'to spring, gush'; at any rate, an old word (Vedic: RV-Khila, AV in Paippalāda version); not as Srinivasan maintains,"only late in Sanskrit, latter than its common use in Prakrit texts". --- Perhaps IE.

- Sum. zi 'life' :: Beng. ji, jio 'to live'.

Ved. *jīvati* 'lives' > Pāli *jīvati*, Pkt. *jīvaï, jīaï, jiaï* > NIA: Nep. *jiunu*, Beng *jiyā*, Hindi *jīnā* CDIAL 4241, cf. 5239-5252; EWA I 594 < IE *$g^w ih_3$-*we-*; Engl. *quick, to quicken*, Latin *vīvus*, Slav. *živ*, etc. --- Transparent IE derivation.

"Three-consonantal stems"

- Sum. da-gal 'wide, broad' :: Beng. dāgar 'large, big, broad'.

Necessitates a reconstruction such as MIA **daggara* < OIA **dargara*, etc., which renders it different from Sum. --- Unclear.

- Sum. za-lag 'shining, bright' :: Beng. jhalak 'shining, flashing, be bright'.

Skt. *jhala* 'flash', Skt. lex. *jhallikā* 'light', lex. *jhalā* 'blaze of the sun', Pkt. *jhalā* 'mirage', cf. Pkt. *jhalajhalaṃta* 'shining' > Nep. *jhaljhal* 'bright, ablaze', Hindi *jhaljhalānā* 'to shine' CDIAL 5352, EWA III 216 "unclear". --- Unlear, cf. below on 'onomatopoetic' words.

- Sum. su-húr 'braids, chignon' :: Beng. cikur 'hair, hairdo'.

Tatsama from Class. Skt. *cikura* 'hair of the head'; normal development > Pkt. *ciura, cihura* 'hair of the head' > NIA: Hindi *cihur* 'hair of head', etc., CDIAL 4776; EWA III 186 "unclear"

- Sum. nin-da 'bread, food' :: Beng. rān-dhā 'cooked food'. --- Unclear.
- Sum. nin-da 'councel' :: Beng. nindā 'criticism, saying bad things about a person'.

Ved. *nindati* 'blames' > Pāli *nindati*, Pkt. *niṃdaï* > NIA: Panj. *nindṇā*, Marathi *nīnẽ*, etc.; cf. Ved. *nindā* 'blame', etc. CDIAL 7211; EWA II 54 (s.v. *ned*), Avest. *nāismī* 'I denounce', IE **$h_3$neid*. --- Transparent IE origin.

- Sum. nagar 'carpenter' :: Beng. lagur 'wooden staff, timber'.

From early Pkt.: Late Ved., Epic *lakuṭa* 'club, stick' > Pāli *lakuṭa, laguḍa*, Pkt. *laguḍa, laüḍa* > NIA: Panj. *lauṛā*, Nep. *lauro* 'stick', Hindi *lauṛā* 'penis'; cf. Pkt. *lakkuḍa* 'stick', Panj. *lakkaṛ*, Hindi *lakkaṛ, lākaṛ*, Mar. *lākuḍ* 'wood' CDIAL 10875; EWA II 472 "unclear', (cf. EWA II 406 on: **laṣṭi*, CDIAL 10991, *yaṣṭi* 'stick' 10444, KEWA III 84). --- Note also that Sum. *nagar* is supposed to be from a local Mesopotamian, pre-Sumerian substrate! --- Separate local substrate word in Sum. and Beng.

- Sum. ka-lal 'honey-mouth' used for sweet words :: Beng. kallol 'sweet sounds, bird's songs'.

Class. Skt. *kalanā* 'inciting', Pkt. *kalaṇā*, Beng. *kalnā*, *kallā* 'coquetry', CDIAL 2916, EWA III 70 "unclear". Probable origin in DEDR 1302 Tam. *kala-kala* 'to reiterate in sound", cf. also DEDR 1302 Tam. *kallal* 'disturbance, tumult, noise due to many people speaking at the same time', DEDR 1310 Tam. *kalipali, kalipili* 'uproar, disturbance, quarrel', Tel. *galibili*, Mar. *galbal-nẽ* 'to be in commotion'; note Mundari *kal-kalāo* 'to make a noise'. --- Perhaps loan from Drav.

- Sum. kik-kin 'fine milled' :: Beng. cikan 'very fine' (adj.).

Epic, Class. Skt. *cikkaṇa* 'slippery, unctuous', Pkt. *cikkaṇa* 'sticky, oily' > NIA: Beng. *cikan* 'smooth, glossy, fine, thin', Maithili *cikan* 'slippery, polished', Hindi *ciknā* 'oily, fat, rich, smooth', etc., CDIAL 4782. Tatsama from Skt./Pkt. - Perhaps from Drav.: Telugu *jiguru* 'gum, birdlime', *cikkana* 'thickness of a liquid', etc., DEDR 2488, but note also Munda: Sant. *cikāṛ, cikan, cikon, cokon* 'smooth, glossy, oily, polished' which cannot all be explained as loans from Hindi or Bengali. --- Loan word from Drav.?

- Sum. ki-kal-la 'hard and dry ground, barren places' :: Beng. kākaṛ 'tiny piece of stone', kākuṛe 'land or soil, means the same as ki-kal-la'.

Skt. (lex. only) *karkara* 'stone, hard' > Pkt. *kakkara*, 'stone, pebble' > NIA: Beng. *kā̃kar* 'gravel', Hindi *kā̃kar* 'nodule of lime stone', etc.CDIAL 2820, cf. also 2810 *karkara* 'hard', 2822 *karkaśa* 'rough, hard'; EWA III 65 deliberates on a vague IE and on local origin: Skt. *karkara, karkaśa, kakkhaṭa* 'hard', *khara* 'rough', which points to several separate takeovers of local loan words; cf. also MT SI 1999 *śarkara* 'sand, pebbles' and Bur. *ɣoro* 'stones'.

- Sum. kan-kal 'waste land' :: Beng. kañkar 'rubbles', used for barren places.

See preceding : CDIAL 2820 Skt. (lex. only) *karkara* 'stone', Pkt. *kakkara*, 'stone, pebble', Beng. *kā̃kar* 'gravel', Hindi *kā̃kar* 'nodule of lime stone', etc. --- Local loan word.

- Sum. za-dim 'jeweler', za 'precious stones' :: Beng. jaṛoā 'jewelry with precious stones', jaruā (M.Beng.) 'jeweler'.

OIA **jaḍati* 'joins, sets', Pkt. *jaḍia* 'set (of jewels), joined > NIA: Kashm. *jarun* 'to set jewels', Beng. *jarāna* 'to set jewels, wrap around, entangle', Old Awadhi *jaraï* 'sets jewels, bedecks', etc. CDIAL 5091. --- Local South Asian *(-ḍ-)* loan word.

- Sum. bar-rim-(ma) 'dry land' :: Beng. bārind 'high and dry land'.

Skt. (class., lex.) *varaṇḍa* 'mass, heap of grass, grassy knoll', *varaṇḍikā* 'small mound of earth' (note that even a small elevation in most of W. Bengal and Bangla Desh means the difference between wet and dry land), Pkt. *varaṃḍa* 'wall' > NIA: Nep. *barandā* 'raised platform before the house', Or. *araṇḍa* (sic) 'veranda', CDIAL 11317; EWA III 456 "unclear". --- Local substrate word, with typical S. Asian -*ṇḍ*-.

- Sum. ba-ra 'negative verbal prefix' :: Beng. bārān prohibition, 'not to do'.

Tatsama = Epic Skt. *vāraṇa* 'act of obstructing' > Pāli, Pkt. *vārāṇa* 'obstruction, obstacle' > NIA: Or. *bāraṇa* 'obstruction, dam' CDIAL 11553; all from Ved. *vārayate* 'obstructs, keeps back' > Hindi *barnā* 'to prohibit, leave off, separate' etc., Beng. *bārā* 'to shield, stop', Or. *bāribā* 'to obstruct, prevent, forbid', etc. CDIAL 11554; EWA II 512 from IE **wel* 'to encompass, cover' etc. --- Transparent IE origin.

Onomatopoetics

The origin of the words called 'onomatopoetic' in Srinivasan's lists are, as always, much more difficult to trace, as they are not always covered by the standard dictionaries. First, we need a proper definition of the term. South (and S.E., E.Asian) onomatopoetics, also called ideophones or expressives, cover a large range of sensual impressions from hearing, touch and sight to some rather abstract notions. A good definition and discussion of South Asian onomatopoetics has been given by Pinnow 1959: 19 sqq., for Drav. by Emeneau 1969, for Skt. by K. Hoffmann 1962, and more recently by Abbi 1987 and Masica 1991: 78-81).

Pinnow 1959:19 defines them (including echo words) as providing the 'designation of variable objective indications about size, weight, form, temperature, brightness, variation of color; further, indication of positive and negative value judgments, imitation of sound and movements of all sorts; finally, intentionally or unintentionally imprecise, intensifying or weakening, praising or denigrating expressions' (my translation).

There are, in my opinion, several types of S.Asian onomatopoetics, starting with (1) IE types (cf. Greek *bar-bar-oi*, Engl. *murmur, sing-song, splish-splash*): Ved. *bal-bal* 'sound of bottle, leather pipe emptying' (still the same, *balbal*, in Hindi), added to by the later Vedic, "Dravidian" type *kala-kala*, e.g. Epic Skt. *hala-halā* 'applause'; DEDR 1162 *kaṇakaṇa* > Skt. CDIAL 3791 *khaṇakhaṇayate*, etc.

(2) The Drav. type is indeed still common in modern Tamil etc.: *kala-kala* [kala-xala] 'to reiterate in sound, rustle, tinkle, chink, clink, rattle" DEDR1302, *Tam. kaṟa-kaṟa* 'to crackle in the mouth (as a crisp cake)' DEDR 1386, or Kan. *gara-gari-kle* 'pleasantness', DEDR 1259.

(3) the Munda type: substitution of the initial consonant in the second, repeated word by another sound that usually preserves the point of articulation, but allows interchange of velar with retroflex/dental. There are innumerable variations, e.g. "*pata, pada, patha, pana, panda, banda, bada, badha, bana, mata, mana, manda* etc.," as in the Santali word for 'lame': *kəḍuə, kədu'j,... kaḍa', koḍo',... kəḍgu'j, kaṛa'ḍ, kore'ḍ,... khordha, khoṛdhe'ḍ,... huṭe'ḍ, huṛteng*, etc. (Pinnow 1959). Finally, cf. also Burushaski*: ḍaḍanuc ḍupḍup* 'drumming', *guḍun* 'knocking', etc. (Berger 1998 : 215).

One must also take note of some special types, in RV: *bababā* 'sound of fire,' *jañ-jañ-ā-* (bhavat) 'flickering, flaming'; *karkari* 'lute' (with similar words in Drav., Munda: Sant. *karkur, gaṛgaṛ, gargor* etc.) - Note that the Sumerian examples, below, still have another type, more like the RV *bababā: tál-tál-la, dag-dag-ga*, if we may take the written representation at face value here.

Then, there also are the so-called echo words (Masica 1991, Pinnow 1959), with various substitutions of the initial consonant in the repeated word, as seen in Hindi *roṭī-śoṭī* 'all kinds of bread' (*roṭī*), *kitāb-mitāb* 'various kinds of books' (loan from Arab. *kitāb*!). Drav. and Munda have the same doubling and change of initial consonant, Tam. *kalipali, kalipili* (below), Sant. *ṭakar bakar*.

Nowadays, and indeed since Vedic times, these types are no longer found in splendid isolation, but the various forms have been exchanged all over S. Asia, so that it is difficult to establish which form is original in a particular case, i.e. IE type in Drav.: Kurukh *xar-xar* 'perfectly clean' DEDR 1259, or the Munda type in Tam. *kalipali, kalipili* 'uproar, disturbance, quarrel', Tel. *galibili*, etc.

Another important feature of these words is that they frequently are not affected by phonetical changes, for example, Jpn. *peko-peko* 'hungry' (sound of empty stomach) or *pata-pata* 'fluttering' should not have initial *p-* as this has changed to *h-* some 1500 years ago; secondly, the repeated word should have a Sandhi form, **peko-beko* > mod. Jpn. **heko-beko*, **pata-bata* > mod. **hata-bata*, which is not the case (and probably would not have conveyed the intended 'expressive' meaning). In the same way, Vedic (Jaiminīya Brāhmaṇa) *bal-bal* is still retained in modern Hindi (differently, Masica, 1991).

Finally, reduplication, whether 'onomatopoetic' or not, is of course every common in many languages of the world to denote repetition of action or plurals, from French (child speech) *bon-bon* 'sweets', to Indonesian *orang-orang* 'men, people' or Jpn. *hito-bito* 'men, people'.

For all these reasons, most of the following words are difficult to etymologize and, often enough, they are not found in the dictionaries.

- Sum. zubi 'water logged ground' :: Beng. job-job 'very wet, watery'.
- Sum. gal 'large, big, great' :: Beng. gal gal 'excessive, much'.
- Sum. dar 'continuous flow' :: Beng. dar dar 'continuous flow of tears, blood'.

Cf. similar 'onomatopoetica': Jpn. *tara-tara* 'dropping, dribbling,' the Engl. sound imitation *d-r-* in *drop, drip*, and contrast IE **drā, dre-b, dre-m, dre-w* 'to run, walk in small steps' etc. Pokorny p. 204; clearly, the pre-PIE root was ***der, dṛ*. We have to distinguish three levels in such IE 'expressive' words: 'original' (pre-)PIE onomatopoetic, PIE verbs and nouns built on such roots, and their later (Engl., etc.) derivations. Quite apart from all of this, there is *retention* of sound symbolism, often in unchanged form: note above, regular development PIE **dr* > Engl. *tr-* (but not *dr-*!) as found in PIE **denk'*, Skt. *daṃś*, Greek *dánk* 'to bite' :: Engl. *tongs* 'pincers'; conversely, Skt. *tata* 'father', which is kept in various IE languages, in spite of expected sound changes, cf. also Turkish *ata*, etc., see Pokorny 1056. -- Finally, note the expected **dr-* > *tr-* in Engl. *tr-* in *to trip*!

- Sum. tuku-tuku 'tremble' :: Beng. ṭhak-ṭhak 'accompanies verb to tremble' as in bhoe ṭhak-ṭhak 'trembling in fear'.

OIA **ṭhakk* 'clatter' (onom.), Nep. *ṭhakṭhakyāunu'* to knock against', *ṭhakṭhuk* 'thumping and bumping', Hindi, Guj. Mar. *ṭhakṭhak*, etc. CDIAL 5487.

- Sum. dag / dadag , dag-dag-ga 'washed, to be clean' :: Beng. dag-dage/ dag-dag 'to be perfectly clean'. Sound of beating clothes when washing; -- or, as Mark Twain has it about Indian washer women, 'to split a stone'. Onomat., as in Engl. *tick-tack*.
- Sum. tál, tál-tál-la 'to make wide, wide spreading' :: Beng. ḍhal ḍhal 'too wide i.e. loose fitting clothes'.

Cf. CDIAL 5581 *ḍhalati* 'bends over, falls'??

- Sum. sir-sir 'serpents' :: Beng. sar-sar 'qualifying movements of reptilian creatures'.

Old onomatopoetic? Note: PIE **ser* 'to stream, to move to and fro quickly', Pokorny 909 :: PIE **sṛp*, Ved. *sarpati* 'to creep', *sarpa* 'snake' > Nep., Beng., Hindi, etc. *sāp* 'snake', CDIAL 13271.

- Sum. gaz 'to cut off' :: Beng. ghac-ghac accompanies verb 'to cut' . Unclear.
- Sum. ti-en, ti-en ti-en 'to be cool, cold' :: Beng. ṭhan ṭhan 'cool' (in place names), ṭhan-ḍa 'cold.'

Ved. *stabdha* 'firm, fixed' > Pāli *thaddha* 'slow', Pkt. *ṭhaḍḍha* 'blunted, cold', Nep. *ṭhaṇḍā*, Hindi *ṭhaṇḍ(ā)* 'cold', etc. for the development of meaning, 'firm' > sluggish, cold' , see CDIAL 13676, EWA II 754 *stambh*, past participle *stabdha*, from IE **stembh* 'to prop up, stiffen' --- Inner-Indic semantic development from transparent IE word.

- Sum. mur 'to crush, to grind' :: Beng. maṛ maṛ, muṛ muṛ, sound of crushing, grinding.'

Cf. CDIAL 10187 *muṛ* < 'defective'; Skt. *mṛ, mṛṇāti* 'grinds, crushes'; but note already in MIA > *maṇāti* 'crushes' > NIA Kashm. *munun*. Note the many forms in PIE (**mel, mel-d, mel-dh, ml-ei,* etc.) based on ***mel* 'to grind, crush', Pokorny 716.

- Sum. zal-zale 'bright, brilliant' :: Beng. jal-jale 'to be bright, shining or glittering'. See above, Skt. *jhal* 'flash' > NIA: Beng. *jhalak* 'shining, flashing, be bright', CDIAL 5352, EWA III 216.

- Sum. za-pa-ág 'noise, sound' :: Beng. jhapāg 'noise, sound of something falling'. OIA **jhapp* 'sudden movement', Class. Skt. *jhampa* 'jump' > NIA: Or. *jhap* 'hastily', Beng. *jhap* 'splash'? CDIAL 5336.

Summing up

As in her earlier paper on Canaanite words (MT I), L. Srinivasan assumes western sub- or adstrates in North-Indian languages, especially in Bengali. From a review of the words given above it is clear, at least to me, that most of them can be explained from older, *actually existing* forms of the same words within older Indo-Aryan itself (and a quite few from other (in part lost) languages such as Dravidian, a task carried out only sporadically here). These older forms do not have any close resemblance to Sumerian.

There are, for the time being the following remnants. Note that most words with internal (or initial!) retroflexes, especially *-ṇḍ-*, should be excluded from the list of possible candidates for Sumerian comparison, as this is a typical S. Asian feature encountered from the beginning of our textual sources (Ṛgveda, c. 1500 BCE) onwards; it is not found in the closely related Mitanni IA in W. Asia.

CULTURAL LOANS: 'WANDERWÖRTER'

- Sum. *gi-gir* 'wagon' :: Beng. *gāṛi*
- Sum. *kemu* 'flour' :: Beng. *gam* 'wheat'
- Sum. *uru* 'city, town' :: Beng. *ūr*, Pkt. *ūra*

EARLY IIr LOAN WORDS IN CENTRAL ASIA

- Sum. *úr* 'thigh' :: Beng. *ūru* 'thigh'
- Sum. *ku* 'to eat' :: Beng. *khā* 'to eat'

LOCAL LOAN WORDS FROM S. ASIAN SUBSTRATES

(note retroflexes!), some maybe indirectly from Sumerian (23)

- Sum. *giš* 'tree' :: Beng. *gāch* 'tree' < **gakṣa* (NW *-kṣ-*)
- Sum. *giri* 'foot' :: Beng. *goṛ* 'foot'
- Sum. *gi-gir* 'wagon' :: Beng. *gāṛi*
- Sum. *udu* 'sheep' :: Beng. *udo* 'sheep, stupid'
- Sum. *gar* 'to do, make' :: Beng. *gaṛā* 'to do, make'

- Sum. *gúr* 'to bend down, bow' :: Beng. *gaṛ* 'to bend down to show respect' <*gaḍ(h)-
- Sum. *gúr* 'to twist, roll up' :: Beng. *gaṛā, gaṛo* 'to roll'
- Sum. *pad* 'to tear out, remove' :: Beng. *pāḍ* 'to tear out'

Sum. *rar(r)* 'to cut', dar 'to split' :: Beng. *thoṛ, thoṛā* 'to chop, rend'
- Sum. *tar(r)* 'to break' :: Beng. *toṛ, toṛā* 'to break'

?• Sum. *guz* 'to crush, smite' :: Beng. *ghuc, ghucā* 'end, destroy with force or violence'

- Sum. *kuš* 'skin, hide, leather' :: Beng. *khosā* 'skin (of vegetables)', *khos* 'scab, skin disease', *khuski* 'dandruff'.
- Sum. *gun* 'tribute, talent' :: Beng. *guna-gār* 'to pay/compensate money'
- Sum. *kur* 'mountain' :: Beng. *cuṛ, cuṛā* 'peak of mountain'
- Sum. *sur* 'to squeeze, to press out ' :: Beng. *šuṛi* 'the caste name of the wine maker and seller'.
- Sum. *mun* 'salt' :: Beng. *nun* 'salt'

- Sum. *na-ru-a* 'stela' :: Beng. *noṛā* 'term for round and elongated grinding stone'
- Sum. *nagar* 'carpenter' :: Beng. *lagur* 'wooden staff, timber'
- Sum. *ki-kal-la* 'hard and dry ground, barren places' :: Beng. *kākaṛ* 'tiny piece of stone', *kākuṛe* 'land or soil, means the same as *ki-kal-la'*
- Sum. *kan-kal* 'waste land' :: Beng. *kañkar* 'rubbles', used for barren places
- Sum. *za-dim* 'jeweler', *za* 'precious stones' :: Beng. *jaṛoā* 'jewelry with precious stones',

- Sum. *bar-rim-(ma)* 'dry land' :: Beng. *bārind* 'high and dry land';
- Sum. *kar* 'town' :: Beng. *gaṛ* 'fortified town'
- Sum. *šu* 'hand; to do things with the hand' :: Deśī *ço, çu (śo, śu)* 'to touch, feel with the hands'; *śur* 'tentacles, trunk (used like hands)'

REMNANT, SO FAR UNEXPLAINED WORDS (17)

- Sum. *śá-an* 'clever' :: Beng. *seyanā, seanā, seyān* 'clever'
- Sum. *nitah* 'male' :: Beng. *nitā* 'male worker'
- Sum. *tum* 'abundance, plenty' :: Beng. *dhum* 'plentiful, abundant'
- Sum. *til, til-til* 'to pick off, pluck' :: Beng. *tolā* 'to pick off, pluck'
- Sum. *gur* 'to run about' :: Beng. *ghurā* 'to travel, to travel around in a place'

- Sum. *gur-gur* 'to reel' :: Beng. *ghurāṛ, ghorāna* 'to reel'
- Sum. *gur* 'to return, come back' :: Beng. *ghur* (with aux.) 'to return, come back, give back'

?• Sum. *gir/ger* 'to enclose, to grid' :: Beng. *ghirā/gherā* 'to enclose, to grid, to surround'

- Sum. *sila* 'measure of capacity' :: Beng. *sera* 'a traditional unit of capacity'
- Sum. *iku* unit of area' :: Beng. *ekar* 'traditional unit of area'

- Sum. *tur* 'to be small in size, weight', *tur-ra* 'small, childish' :: Beng. *thorā* 'little, small'
- Sum. *ma-lah/ má-lal-a* 'boat man, sailor' :: Beng. *mālā, mallā* 'caste of mariners, sailors, fishermen'.
- Sum. $^{\mathit{GI\check{S}}}$*má-šu* (-a) 'ship, deep draught boat' :: Beng. *mājhi* 'a caste of boat rowers and boatmen'
- Sum. *da-gal* 'wide, broad' :: Beng. *dāgar* 'large, big, broad'
- Sum. *su-húr* 'braids, chignon' :: Beng. *cikur* 'hair, hairdo'

- Sum. *nin-da* 'bread, food' :: Beng. *rān-dhā* 'cooked food
- Sum. *dingir* 'god' :: Beng. *dāngar sāñi* 'a local god, worshipped by boatmen'

'ONOMATOPOETIC WORDS' (11)

- Sum. *gal* 'large, big, great' :: Beng. *gal gal* 'excessive, much'
- Sum. *dar* 'continuous flow' :: Beng. *dar dar* 'continuous flow of tears, blood'
- Sum. *tuku-tuku* 'tremble' :: Beng. *ṭhak-ṭhak* accompanies 'verb to tremble' as in *bhoe ṭhak-ṭhak* 'trembling in fear'
- Sum. *dag / dadag , dag-dag-ga* 'washed, to be clean' :: Beng. *dag-dage/ dag-dag* 'to be perfectly clean'
- Sum. *tál, tál-tál-la* 'to make wide, wide spreading' :: Beng. *ḍhal ḍhal* 'too wide i.e. loose fitting clothes'

- Sum. *gaz* 'to cut off' :: Beng. *ghac-ghac* accompanies verb 'to cut'
- Sum. *mur* 'to crush, to grind' :: Beng. *maṛ maṛ, muṛ muṛ,* sound of crushing, grinding'
- Sum. *za-pa-ág* 'noise, sound' :: Beng. *jhapāg* 'noise, sound of something falling'. --- cf. also earlier in the paper.
- Sum. *i-lu* 'song, wail' :: Beng. *ulu* 'ceremonial sound'
- Sum. *pá(d)* 'to call, to speak' :: Beng. *pāḍ* in *dāk pāḍ* 'call', *hā pāḍ* 'to shout'
- Sum. *dug* 'jug, pot' :: Beng. *ḍhāk* 'large drum', *dugi-dugi* 'small drum'

Some additional remarks

A few stray remarks may be added about some other aspects of the paper. The famous Indian linguist S. K. Chatterjee has compared Sumerian *pilaqqu* 'axe' and *apsu* 'water' with the RV words *paraśu* 'axe' and *ap* 'water'. Actually, Sum. *pilaqqu*, Akkad. *pilaqqu, pilakku* means 'spindle, dagger' (EWA II 87), and Ved. *paraśú*, S.W. Old Iran. **paraϑu* > Ossetic *færœt*, rather are to be compared to Greek *pélekus* (< IE **pelek'u-*) whose exact locus of origin remains unclear (cf. KEWA II 213, and MT, SI 1999).

The Sum. primordial deity of sweet waters underneath the earth, *Apsu*, would only correspond to the Skt. loc. pl. of water, *ap-su*. Skt. *ap-* is derived from a more eastern IE form $*h_2ep$- (Tocharian *āp-*, Hittite *ḫap-*, Baltic *ap-*) that corresponds to more western $*h_2k^w$ (Latin *aqua*, Germanic: *ache, aa*, etc.). Both comparisons, constantly quoted in India, thus are highly unlikely.

Drav. *uru* 'town' is *not* present in the Ṛgveda, as stated by Srinivasan. Rather, the common word *uru-* means 'wide, broad' and is related to the equally common Avestan *vouru* < IE **wr̥H-u-*, Greek *eurú-s* < **werHu-*. In the Ṛgveda, it is used with words signifying wide, open spaces (for cattle); the name *Uru-kṣaya* means 'widely governing', *Uru-kakṣa* 'having wide bushes/brushwood', cf. the name of a poet *Kakṣī-vant* 'having bushes' (cf. above on *gacch*, EWA I 288). No towns here, as they are indeed not found in the RV otherwise.

Origin of some Skt. words

The following (old) Skt. words have been proposed for Sumerian origin. Yet, they have clear IE antecedents.

- Sum. *tag* 'to leave, abandon' :: Skt. *tyāga.*

Ved. *tyaj* 'to leave, give up', Ved. *tyāga* 'the act of giving up, offering' > MIA: Pāli *cāga*, Pkt. *cāya*, O.Singh. *caya* 'gift', otherwise lost. Also, O.Avestan *iϑiiejah* 'loneliness' < IE **tyeg*^{w} 'to give up, withdraw', Greek *sébomai, sobéō* EWA I 673-4. The initial cluster **ty-* is singular, but undeniable; thus not ~ Sum. *tag*. --- Obvious IIr, IE root.

- Sum. *tab* 'to burn, to heat up' :: Skt. tāpa 'heat', Actually Ved. *tapas* 'heat', *tapuṣ, tapnu* etc., from *tap* 'to heat', Avest. *tāpeiieti* 'heats up' < IE **tep* 'to be warm', Latin *tepēre* 'to be warm', etc. EWA I 623-4. At best, long range connection between IE **tep*, Sum. *tab.*
- Sum. *dam* 'wife or husband' :: Skt. *dam-pati* 'wife and husband, married couple'.

This actually is *dam-patī* 'wife and husband' (dual of *dam-pati-* 'lord of the house') = 'the two lords of the house, husband and wife'. From Ved. *pati* 'lord, husband' < IE **poti* (Latin *potis* 'able, can do', Greek *pósis*) + *dam* < IE **dem* 'house' (Ved., Avest. *dam-* 'house', Latin *domus* 'house'); this is already an IE formation: **dém-s pótis*, EWA II 73-4; no connection with Sumerian.

Finally, L. Srinivasan's closing remarks about the "urgent need to reexamine the vernaculars, and ancient Indian texts anew (the current available interpretations are 100 to 150 years old)" remain true, except for the fact that the available etymological dictionaries (CDIAL, EWA, Pokorny, all not used by Srinivasan) are up to date, and that textual interpretations are proceeding apace, as any bibliography of Vedic literature would indicate (Renou 1931, covering everything written up to that date, and Dandekar 1946sqq. for all that has appeared until some ten years ago).

What we *really* need, are detailed etymological studies, but not of folk etymology (Nirukta) kind,[110] studies that are virtually absent in India: there are no etymological dictionary, exept for Bengali (by S. Sen, 1971, quoted by Srinivasan but not available to me), and CDIAL, EWA, DEDR... This is a wide open field: we would first have to eliminate, from the NIA languages, all non-IA words introduced from Persian and Arabic in the second millennium; only then, we can proceed, in collaboration with specialists of Austro-Asiatic, Tibeto-Burmese, Burushaski and Dravidian to a better understanding of the history of South Asian languages. Nothing of the sort is in sight.

---

[110] Just as Latin (*lucus a non lucendo!*) or Greek (Plato's Kratylos), Sanskrit has an old tradition of etymologies of the folk etymology type, such as Ved. *aśva* 'horse' from *aśru* 'tear', with the appropriate mythological background supplied alreday in the Vedas. This tradition continues unabated until today, with cases such as Assyrians = *Asura* (demons), Phoenicians = *Paṇi* (a group of mythological cow thieves), *Sura* (gods) = Syrians, *Maitrāyaṇīya* (a Veda recension and their transmitters) = Mitanni, *Kaṭha* (another Veda recension) = Hittites (Khet), etc.; cf. above on Mesopot. *Apsu* = Ved. *ap-su*, or AV *Taimata* = Mesop. *Tiamat.*

ABBREVIATIONS

Akkad.	Akkadian
Avest.	Avestan
AV	Atharvaveda
Beng.	Bengali
CDIAL	see Turner
Drav.	Dravidian
DEDR	see Burrow and Emeneau
EWA	see Mayrhofer 1951-1975
IA	Indo-Aryan
IE	Indo-European
IIr	Indo-Iranian
Iran.	Iranian
Jpn.	Japanese
Kan.	Kannada
Kashm.	Kashmiri
KEWA	see Mayrhofer 1986-
lex.	lexicographers
M.	Middle
Mal.	Malayalam
Mar.	Marathi
MIA	Middle Indo-Aryan
MT, SI	Mother Tongue, Special Issue
N.	New
Nep.	Nepali
NIA	New Indo-Aryan
Nur.	Nuristani (Kafiri)
O.	Old
Or.	Oriya
Pers.	Persian
OIA	Old Indo-Aryan
Panj.	Panjabi
PIE	Proto-Indo-European
Pkt.	Prakṛt
RV	Ṛgveda
Singh.	Singhalese
Skt.	Sanskrit
Slav.	Slavic
Sum.	Sumerian
Tam.	Tamil
Tel.	Telugu
Ved.	Vedic

REFERENCES

Abbi, A. *Reduplicative structures in South Asian languages: a phenomenon of linguistic area.* New Delhi: Jawaharlal Nehru University; Centre of Linguistics and English. 1987.

Berger, H. *Die Burushaski-Sprache von Hunza und Nager.* Wiesbaden: Harrassowitz 1998.

Blažek, V. and C. Boisson. The Diffusion of Agricultural Terms from Mesopotamia. *Archív Orientalní* 60, 1992, 16-37

Burrow, T. and M.B. Emeneau. *A Dravidian Etymological Dictionary.* Second Edition. Oxford: Clarendon Press 1984

Dandekar, R.N. *Vedic Bibliography; an up-to-date, comprehensive, and analytically arranged register of all important work done since 1930 in the field of the Veda and allied antiquities inlcuding Indus Valley civilisation.* Bombay: Karnatak Pub. House, (vols. 1-4) 1946, Poona 1962, 1973, 1984.

Emeneau, M.B. Onomatopoetics in the Indian Linguistic Area, *Language* 45, 1969, 274-299

Hoffmann K. Rezension von Wackernagel, Altindische Grammatik. *Indogermanische Forschungen* 60, 254-264 = *Aufsätze zur Indoiranistik*, ed. J. Narten, vols.1-2. Wiesbaden. 1975-76

Kuiper, F.B.J. *Proto-Munda words in Sanskrit.* Amsterdam: Noord-Hollandsche Uitgevers Maatschappij 1948
Aryans in the Rigveda. Amsterdam-Atlanta: Rodopi 1991

Mayrhofer M. *Kurzgefasstes etymologisches Wörterbuch des Altindischen.* Heidelberg 1956-1976. (KEWA)
---, *Wörterbuch des Altindoarischen.* Vol. I-II [Vedic] Heidelberg 1986-96, Vol. III [Epic, Class. Skt.] 1997sqq (EWA)

Nirukta: *The Nighantu and the Nirukta : the oldest Indian treatise on etymology, philology, and semantics / critically edited from original manuscripts and translated for the first time into English, with introduction, exegetical and critical notes, three indexes and eight appendices by Lakshman Sarup.* London/New York: Oxford University Press 1920-1929.

Pinnow, H.-J. *Versuch einer historischen Lautlehre der Kharia-Sprache.* Wiesbaden 1959

Pokorny, J. *Indogermanisches etymologisches Wörterbuch.* Bern/München 1959

Masica, C. P. *The Indo-Aryan Languages.* Cambridge: Cambidge University Press 1991

Renou, L. *Bibliographie védique.* Paris: Adrien-Maisonneuve 1931

Sen, S. *An Etymological Dictionary of Bengali: c. 1000-1800 A.D.* Calcutta: Eastern Publishers 1971

Srinivasan, L. and C. Gordon, Canaanite Vocabulary in Bengali and Some Other IE Dialects of India. MT I, 1995, 202-206

Turner, R. L. *A Comparative Dictionary of the Indo-Aryan Languages. London: Oxford* University Press 1966-69

Witzel, M. Early Sources for South Asian Substrate Languages. *Mother Tongue*, Special Issue, October 1999

More on Possible Linguistic Connections of the Sumerians

By Igor M. Diakonoff
St. Petersburg, Russia

[Editor's Note: The following article was sent by **ASLIP Council** Fellow Igor M. Diakonoff to former MT Editor Harold C. Fleming early in 1999. Since this was only shortly before Diakonoff's death, this has, regrettably, become the esteemed scholar's last word for *Mother Tongue.*]

In *Mother Tongue* III (December 1997, **p. 54**) **I published** an article "External Connections of the Sumerian Language." It was followed by two articles on kindred topics: John D. Bengtson, "The Riddle of Sumerian: A Dene-Caucasic Language?" (p. 63), and Allan A. Bomhard, "On the Origin of Sumerian" (p. 75).

I

In a discussion of the sort which started between Bengtson, Bomhard, and myself, one should, to my mind, make it clear to the readers, what the exact topic is. Do we investigate linguistic kinship, or historic connections, or a group of languages connected by common history, perhaps, only by a number of borrowings?

Being a historian as much as (or more than) a linguist, I raised the question where the Sumerians came from, because the territory that they inhabited in historical times in southern Iraq could not have been occupied earlier than from about 3000 BCE, since, at least until late in the IVth millennium, it was completely under water.

The Sumerian legend brought them from Dilmun, i.e., from Bahrain, an island, mostly barren, in the Persian Gulf.

Obviously, rocky Dilmun could not have been the primary place of origin of such an important people as the Sumerians, so we should look for their initial habitat farther to the east of Bahrain.[1] Thus we arrive in southern India, the oldest known inhabitants of which are the groups of tribes speaking Munda languages.

The next step in our explanation was to find lexical elements common to Sumerian and Munda. Now the Munda languages themselves are not quite obviously similar. For comparison, I have selected the Kherwari branch, its most important dialects being Santali and Mundari. These dialects seem to be the most archaic, and at the same time nearer to each other than the rest of the Munda languages, whose difference from the Kherwari might be due to some neighboring non-Munda linguistic influences.

---

1. The territory to the north of the Persian Gulf was inhabited by Elamo-Dravidian speakers, including the Brahui – not akin to Sumerian or to Munda. The territory to the south of the Gulf was mostly desert.

I have found 34 lexical items which are common to Sumerian and Kherwari. Several of them correspond to items on Swadesh's 100-word list. I concluded my article with the words: "I do not presume to have found the final solution, but I have made a start in looking for the needed answer."

John D. Bengtson is one of the brilliant linguists of our time. He easily operates with languages of all possible families and grades of kinship, but, to my mind, he sometimes neglects the problem of historical probability. In his article, Bengtson practically disregards my Munda evidence, and argues that Sumerian is a Dene-Caucasian language, belonging to a linguistic macro-family supposedly including Basque + (North) Caucasic[2] + Burushic + Sino-Tibetan + Yeniseian + Na-Dene, including Chipewyan (i.e. Ojibwe).[3] This list raises serious doubts, or, at least, requires serious corrections.

Then Bengtson lists some Sumerian glosses which, he suggests, may be connected to lexical elements in various parts of the world:

1. BLOOD: Sumerian *guru, gurun, kurun :* Na-Dene: Chipewyan *–gai* ['white'], *gay-í* 'reddish'. (Why not compare Russian *krov'* 'blood'?)
2. To BREAK: Sumerian *pat'* (sic!)[4] : Burushic *phalt-*.
3. FEMALE BREAST: Sumerian *agán*. Bengtson compares the Yeniseian (Kott) *xanti.* Obviously, he could not discover anything less dissimilar. There is absolutely no evidence of historical contacts between Sumer and Siberia.
4. BURN: Sumerian *tab* compared with Tibetan *thab* 'hot (springs)', Chipewyan *–t'a* 'it is hot, roasted'.[5]
5. COLD: Sumerian *te(n)* 'be cold', *en-tena* 'cold, chill, frost'.[6]
6. DAY, SUN: Sumerian $u_4(d)$ 'day', *utu* 'sun' : Basque *uda* 'summer', Tibetan *od* 'light, shine, brightness'. It seems the word must come from a proto-proto-language, including an area from Spain to Tibet. *Sapienti sat!*
7. BEND, BEND DOWN (figurative for 'die'): Sumerian *gam.* The usual Sumerian word for 'die' is *ug.* Bengtson's comparison to Sumerian *gam* is Tibetan *ãgum* 'to die' (elegant). How does one die elegantly?
8. DRINK: Sumerian *nağ̃* : different Na-Dene languages have *–ná, -naah-, naan, -naŋ*, etc. Hurray! But, regrettably, Bengtson's Na-Dene languages are from America.

* * *

[2] NB: There are *several* (sometimes distantly related) linguistic families in the Caucasus. In Daghestan, both Caucasic and non-Caucasic languages are present. A single Daghestani language does not exist.

[3] [Diakonoff confuses Ojibwe (= Chippewa, an Algonkian language spoken around the Great Lakes of North America) with *Chipewyan,* a Na-Dene (Athabaskan) language spoken in northwestern Canada. Bengtson was citing the latter. Ed.]

[4] A Sumerian phoneme **t'* does not seem to have been proved. Read *pad.* [The form actually cited by Bengtson was *pad.r,* following Thomsen and Boisson. Ed.]

[5] [The words actually cited by Bengtson were Tibetan *thab* 'fireplace', Tlingit *t'ay* 'heat, hot springs', Chipewyan *t'ɛ̀* 'it is hot, to be roasted'. Ed.]

[6] [Diakonoff omitted the proposed Na-Dene cognates, e.g., Chipewyan *–tə̀n* to freeze; ice', Beaver es-*təne* 'ice'. Ed.]

This is what the whole list is like. I may add that Chipewyan (Ojibwe)[7] is in North America, Burushic is in the Himalayas, Basque in Spain, Tibetan – in the heart of Asia. If Bengtson is right, we are beholding disconnected parts of the language of Adam and Eve, or, more realistically, *Pithecanthropus*.

Referring to my own paper, Bengtson (p. 73) makes the following statement: "… it is not that Sumerian and Munda are 'unrelated', but that the relationship is probably indirect, by virtue of the common origin of the Dene-Caucasic and Austric macro-families, and too remote[8] to shed any light on the immediate origin of the Sumerians." Unfortunately, no proof of Austric connections of Sumerian are presented to the reader. No Munda lexemes are quoted, nor is it mentioned, to what linguistic family they belong.[9]

II

The second reaction to my paper is "On the Origin of Sumerian" by Allan R. Bomhard. I must say that I agree with most of his statements. The only point where we disagree is on pp. 84-85, where Bomhard says that "the evidence seems to indicate that Sumerian is related to the Nostratic languages as a group," that is to say, as I understand it, that it is a relative of Proto-Nostratic. The historian in me can hardly believe in the co-existence of a huge conglomeration of men and women all speaking still undivided Proto-Indo-European, or Proto-Na-Dene, or, of course, the still undivided Proto-Nostratic, or whatever, and a tiny long-lived people retaining their own non-Indo-European (or non Na-Dene) language, not included either in Na-Dene or Indo-European, but related to them, and intact for millennia.

I have found 34 words common to Kherwari and Sumerian, and among them at least thirteen Kherwari words which belong to the Swadesh 100-word list. And I still believe this result is sufficient to hypothesize that the Sumerians had migrated to southern Iraq, after it became dry about 3000 BCE, from southern India, where they had contacted the Munda. As a historian, I have the right to ask my linguist colleagues: is there any non-linguistic *evidence* that the Sumerians arrived in Iraq from some other place(s)? The ancient inhabitants of the Caucasus and the Near East are known not to speak Sumerian or any kindred language. Their neighbors to the East, the Elamites and the Brahui, are known to have spoken a language akin not to Sumerian, but to Dravidian; the neighbors of the Sumerians to the West spoke Semitic, or at least Afrasian.

Allan Bomhard publishes a list of 142 lexemes found both in Sumerian and in Proto-Nostratic. Of course, the list shows differences in phonetics, but the kinship is pretty obvious. This proves only that Sumerian is a relative of Nostratic, but this does not disprove that the Sumerians, at a certain period of their history, were neighbors of the

---

7. [See note 3. Ed.]

8. Are Yeniseian and Chipewyan *not* too remote from Basque? [Bengtson meant *taxonomic,* not geographic, remoteness. Ed.]

9. [It is difficult to see what Diakonoff meant here. Bengtson was not proposing a special relationship between Sumerian and Austric, but Diakonoff apparently was. Ed.]

Munda. The Sumerian words which I found to be akin to Munda, in a number which should satisfy Sergei Starostin,[10] are not included in Bomhard's list, for no reasons that I can see. Unless the words quoted by me are not Sumerian, but … Munda, at least originally.

* * *

---

[10] [See, e.g., MT II, p. 121. Ed.]

I would like to quote some additional points to my discussion with Bomhard and Bengtson:

Bomhard:

ak/g 'to make'
gen, ginna 'small, child (?)' = *dumu* 'son, child'
dib (not *dab)* 'to grab'
dug$_4$*-ga* is a participle to *dug*$_4$*/du*$_{11}$ 'to say, speak, sing', etc. (not a separate word)
gú 'forehead', not 'head' (which is *sag*)
kur 'mountain': the (?) is unnecessary
g̃ae 'I' (*mae* in Eme-sal)
sal 'nakedness, vulva'

Bengtson:

ki 'earth' (not *gi*$_1$, *gú* !)
sum (not *mu* !) 'give'
deb is Eme-sal for *dù*$_8$
No. 20 shows no [semantic] connection between Sumerian and Sino-Tibetan
sù 'meat' seems wrong
ka (not *kag/k*) 'mouth'
ér 'water' is unknown to me; some mistake?

No words beginning with **r* are quoted, but cf. *ra* 'beat, kill'; Deimel, *Šumerisch-Akkadisches Glossar,* cites about 50 words beginning with *r-*.

Response to Diakonoff

By John D. Bengtson

First, let me say how difficult it has been to get Sumerologists to discuss the topic of possible genetic relationships of Sumerian. Several of my letters to experts on Sumerian went unanswered. Are they totally uninterested in the topic, or do they wish to keep their cherished language forever unique and mysterious? (I have detected a similar attitude among some Vasconists.) Diakonoff, however, did not fit this pattern. He was intensely interested in the genetic classification of languages, as we can see from his important work in Afrasian (Afro-Asiatic), Hurrian, Urartian, and other Middle Eastern languages.

The question of the genetic relationship of the Sumerian language, and consequently the ethnic connections of the Sumerian people, is of tremendous importance. It should be attacked by a coalition of the best historical linguists and the best Sumerologists and Assyriologists. Why has this not been done? Instead, we have so far had to make do with the rather halting attempts of my colleagues and myself. While Diakonoff made a good effort to make historical sense of the Sumerians, I was disappointed by the linguistic side of his article, as explained below.

If the Sumerian language and Kherwari languages are *genetically* related, then a substantial number of basic roots common to both families should be cited. Diakonoff, as he admitted, listed only 13 Sumerian-Munda comparisons pertaining to the Swadesh 200-word list. (Of my 41 Sumerian – Dene-Caucasian comparisons, 37 were of the latter type.) If, on the other hand, Sumerian and Kherwari were related only *areally,* i.e., in a contact situation (as hinted at by Diakonoff in his recent article: "the Sumerians had migrated to southern Iraq … from southern India, where they had *contacted* the Munda" [emphasis added]), then one should be able to cite a substantial number of cultural words (for example, 'wheat, rice, gold, iron, street, brick, anvil, wineskin') that would surely have been borrowed, one way or the other, in a linguistic contact relationship. But among Diakonoff's Sumerian-Munda comparisons, *none* involves a cultural word on both sides of the equation. (The closest approximations were Sumerian *bùluŋ* 'malt', compared with Santali *buluŋ* 'slime, sweat'; and Sumerian *bùru* 'vault, dome, silo', compared with the basic Kherwari *buru* 'mountain, hill'.) So Diakonoff's comparisons showed *neither* genetic nor areal relatedness, and several were so semantically vague as not to help with either hypothesis ('wish, curse' ~ 'bring, fetch, get'; 'cripple, dwarf' ~ 'rejected, refuse'; 'obsessed' ~ 'idiot').

In his critique of my article, Diakonoff repeatedly makes use of the geographic argument against Dene-Caucasian. How, he implies, can such a widely dispersed string of languages (Basque + Caucasian + Sumerian + Burushaski + Sino-Tibetan + Yeniseian + Na-Dene) possibly be a valid genetic unit? Well, we should know that the distance argument *alone* cannot be taken seriously in the scientific determination of genetic classification. One would never suspect that Icelandic and Bengali were related, were it not for the timely recording of ancient languages such as Sanskrit, Greek, and even Gothic, and the subsequent development of philology and historical linguistics. Edward Sapir was ridiculed for proposing that Yurok and Wiyot, in California, were genetically related to the Algonkian family, but now the Algic family is universally accepted. Many

Nostraticists (and paleo-linguists in general) think that Indo-European and Eskimo-Aleut are the geographic extremes of Nostratic (or Eurasiatic), which is approximately the same geographic span as proposed for Dene-Caucasian.

For the moment, let us leave Na-Dene in abeyance, and just look at some of the striking and basic similarities between Sumerian and Sino-Tibetan (selected from my 1997 article):

HEAD: Sumerian *gú* 'forehead'[1] ~ Tibetan m-*go,* Garo s-*ko* 'head'
I: Sumerian *g̃ae* [ŋae] ~ Tibetan *ŋa* 'I, we', Old Chinese **ŋhā* 'I, we'
(Basque *ni* 'I')
KNEE: Sumerian *dùg, du*$_8$*, du*$_{10}$ ~ Burmese *duh,* etc. (Burush. Hunza *–dúmus*)
LIP: Sumerian *sù* ~ Tibetan m-*čhu* 'lip, beak', etc.
LIVER: Sumerian *kin* ~ Old Chinese **ginʔ* 'kidney', etc. (Burush. *–ken* 'liver')
MOUTH: Sumerian *ka* ~ Tibetan *kha,* etc. (Tlingit *χ'á*)
NAME$_1$: Sumerian *mu* ~ Dimasa *mu,* Bodo *muŋ,* etc. < ***măiŋ**
NAME$_2$: Sumerian *sa*$_4$ 'to name' ~ Old Chinese **seŋ* 'name', etc.
(Basque i-*zen* 'name')
SINEW: Sumerian *sa* ~ Tibetan r-*ca* 'vein, root', etc. (Basque *zain*
'nerve, vein')
WHO: Sumerian *a-ba* ~ Burmese *ba* 'what, which', etc. (Burush. *bo, be*
'what, how')
WIND: Sumerian *líl* ~ Burmese *liy* 'wind', etc. (Basque *luia* 'adverse wind')

Sumerian and Sino-Tibetan apparently share a typological feature, not mentioned in my 1997 article, but suggested by Diakonoff (1997: 57): "Sumerian was certainly a tonal language, which is proved by the very numerous homonyms." The similarity of Sumerian and Sino-Tibetan was recognized as early as 1918, by C.J. Ball. I think he was very perceptive, and this is an avenue of research that should be fully pursued.

I regret that Professor Diakonoff cannot continue this discussion.

---

1. Meaning corrected by Diakonoff in his article.

A DISCUSSION OF CLIMATIC INFLUENCES ON LANGUAGE

Randy Foote

Any consideration of long-range pre-historic language changes and relationships must take into account the variations of climate within the time span considered, as well as genetic changes. Human populations have been impelled to movement by the changes in their ecology -- which is the product of climatic state. In addition, climatic change is not necessarily as gradual as is often assumed, as the kind of gradual change that allows humans to adapt over many generations. Until a few years ago, climate change was generally regarded to be deep background to human prehistory, a slow motion effect -- which would have affected populations only over the span of millennia. But this is not necessarily true. Recent research -- particularly the Greenland ice-core projects -- shows that there have been times within the scope of human prehistory when climate has changed radically in the span of one generation, even in one decade.

No one paper can be comprehensive on this subject. Rather, I hope here to offer, first of all, a broad overview of Late Pleistocene and Holocene climate changes that might be helpful to historical linguists and second, a few examples of recent work relating climate change to linguistic change. This includes two examples of sudden climate change that must have been important to linguistic dispersal -- the Black Sea flood, 7600 years ago, and the deglaciation of the Younger Dryas, 12,500 years ago. These might serve as case-studies for future research and discussion.

Climatic Overview:

I begin with a broad overview of the Upper Paleolithic and Holocene climate, the temporal framework within which language first developed and dispersed. I necessarily focus on a few areas that have been of much interest to linguists. A comprehensive presentation of worldwide paleo-climate change, with detailed physical and vegetation maps, can be found at a fine website developed by Jonathan Adams (then of the Environmental Sciences Division, Oak Ridge National Laboratory), at URL: http://www.esd.ornl.gov/projects/qen/adams1.html

At about 131,000-114,000 years ago occurred the Eemian interglacial, a warm period between glacial epochs, much as the current Holocene is considered to be an interglacial. Climate was similar to that of the present, perhaps even somewhat warmer and moister in some areas. This time-frame is at the limit of what can be accurately dated from the Greenland ice-cores, in that the deeper the layer, the more crushed and deformed it is by the weight of the ice above.

Beginning roughly 115,000 years ago, glaciation began again, with steadily increasing cooling. The change from Holocene-type conditions to full Ice Age conditions took no more than 400 years. Ice sheets grew in the North. Generally, climate cooling created increased aridity both because cooler air evaporated less moisture into the atmosphere, and also because more of the Earth's water became trapped in vast ice sheets. Ocean levels dropped by as much at 500 feet during the depths of glaciation. Forests became savannah, and savannah deteriorated to desert. If Modern Man did indeed first develop in Africa, the arid zone of the enlarged North African and Arabian deserts may have imposed a crucial barrier to human expansion out of Africa until the climate changed again.

71,000 years ago: The peak of this glacial epoch (called the Lower Pleniglacial). This peak coincided with – and was likely caused by – a huge volcanic eruption on Mount Toba in Sumatra that is considered to have been the largest volcanic event of the past 450 million years.

From Ambrose, S., "Late Pleistocene Human Population Bottlenecks: Volcanic Winter, and Differentiation of Modern Humans", Journal of Human Evolution, [1998] 34, 623-651
"…The six year long volcanic winter and 1000-year-long instant Ice Age that followed Mount Toba's eruption may have decimated Modern Man's entire population. Genetic evidence suggests that Human population size fell to about 10,000 adults between 50 and 100 thousand years ago. The survivors from this global catastrophe would have found refuge in isolated tropical pockets, mainly in Equatorial Africa. Populations living in Europe and northern China would have been completely eliminated by the reduction of the summer temperatures by as much as 12 degrees centigrade."

This disastrous cold event may well have created a severe population bottleneck, followed by a rapid expansion of humans from a few refugia. This would have had a profound effect on the distribution of languages, with many branches dying out while a few surviving languages would have spread broadly during repopulation.

From 70,000 ya to 30,000 ya, the climate slowly and intermittently warmed, though not reaching the warmer conditions of the Holocene. There is evidence that there were a number of sudden warming and cooling episodes during this period. This was the time of the expansion of modern humans across the continents of the Old World -- and possibly also into the New World. During most of this period, Australia and New Guinea were joined, and this land-mass was separated from the Indonesian land-mass by the Sunda Strait (50 miles or so), which had to have been crossed in the human occupation of Australia (variously dated between 60,000 and 40,000 years ago) -- the first clear evidence of human use of boats. (e.g.: Noble and Davidson, Human Evolution, Language, and Mind, 1996) This use of boats may well be considered to have been the first indisputable evidence of the level of cognition that would require the use of fully modern language. Also, if the 30,000+ years ago dates of American sites such as Monte Verde II and Meadowcroft are valid, this would have been the time when peoples of a Pacific Rim maritime culture first settled the Americas.

This is the critical time-frame for the first great dispersal of language into its most ancient and highest order branchings, as people moved out of Africa and Southwest Asia toward the north, east and west. The first route of settlement was apparently around the rim of the Indian Ocean, all the way to Southeast Asia, with one route branching toward New Guinea and Australia and another branching northward along the East Asian coast, certainly as far as Japan. Europe was most likely settled later, circa 40,000 years BP, through a separate branching out of the Levant.

Most likely these expansions did not occur as a steady movement; the climate was not changing steadily and gradually, but instead was undergoing "jumps" of warmer and colder states. As further research fine-tunes the dating within this period -- as the dating of the past 20,000 years has recently become more precise -- it could well shed more light on deep paleolinguistics.

At about 30,000 years ago, gradual cooling again began, and the Earth moved into a period of severe glaciation that lasted some 14,000 years. Ice sheets again increased, aridity increased, and ocean levels plummeted. Asia and North America were linked by the glacial ice and tundra of Beringia.

This must have been a period of severe stress for the expanded human populations. The cold and arid climate drastically decreased both food resources and livable land area. Some populations (and their languages) must have disappeared altogether. However, the European Upper Paleolithic, beginning with the Aurignacian, shows that some people also adapted and thrived. This is the beginning of clear evidence of higher symbolic abilities – shown indisputably in the rock art of Europe and other areas. Such symbolic and representational evidence should also be taken as undeniable evidence of fully developed language ability.

At 21,000 - 17,000 years ago came the Glacial Maximum (Upper Pleniglacial). This was a severe period, particularly in the upper latitudes, and the few forests that remained were isolated refugia. Beyond these forest islands in the current Temperate Zone there was mostly desert and semi-desert, with vast areas of ice. The greatest aridity, and least hospitable conditions, occurred near the end of this period. Sea levels were very low, again up to 500 feet below those of today. The Mediterranean was a small dammed brackish body of water; the Sea of Japan was essentially cut off from the oceans. People in mid-latitudes would often have concentrated near the coasts (where artifacts are now submerged), because the oceans moderated the climate in those areas. This period was a likely bottleneck for both populations and languages. It would also have been a period of accelerated linguistic drift between the various isolated and stressed pockets of population.

About 14,000 years ago, warming began rapidly. In many areas, climate seems to have gone from glacial to modern within a few years. The great melt-off began, as the ice-trapped freshwater rushed into great rivers and inland seas. In the Arctic, glaciers retreated from coastal areas soon after the Glacial Maximum, opening what may have been the most likely route for immigration into the Americas from Asia, long before it was possible for hunters to travel south from Beringia by land because of the continuous line of glaciation.

From 14,000 ya to 12,800 ya, there was a continuation of this relatively rapid warming and decrease in aridity. Woodland reappeared in Europe; this caused the decline and the displacement eastward of the Upper Paleolithic large mammal hunting cultures. The Paleolithic big-game hunters began to move north and east, following the remaining herds through Northern Eurasia. The Middle East and Northern Africa were well-watered savannah, optimal for hunter/gather cultures. This was the time of the Great Hunt in the Saharan savannah, evidence of which can be seen in early Saharan rock art. The Sahara and Arabian deserts shrank to less than half their present size. Population expanded because of the increased food supply. Mesolithic, post-glacial culture spread throughout southern Eurasia, as big-game hunting was supplanted by small game and fishing with more reliance on gathering of edible plants.

In North America, the glaciers retreated sufficiently by about 12,000 years ago to open the Cordilleran Gap, an ice-free corridor allowing passage from Alaska to the North American Great Plains. This has been considered to be the path taken by the Clovis hunters, formerly considered to be the first immigrants to the Americas. By 12,000 years ago, the Bering Land Bridge had disappeared.

At the time when the mid-latitude environments improved, the melting glaciers released vast amounts of water. Throughout this period, the ocean levels steadily continued to rise. During the early part of this period, vast freshwater "Great Lakes" appeared in North America and Northern Central Asia, fed by melting glaciers and filling the deep depressions in the Earth's surface left by the weight of the glacial ice-mass. The rivers carried the meltwater southward, and Central Asia must have possessed more water and ecological resources than at any time since the Glacial Maximum. (See below, The Black Sea Flood)

Suddenly, about 12,800 years ago, occurred the Younger Dryas -- an interval when near-glacial conditions of cold and aridity returned suddenly, within 40 years or less. This is a period that has been intensely studied by the Greenland Ice-core Project. Deserts returned to being more extensive than at present. This would have been a time of great population stress for peoples who had begun to expand and adapt to post-glacial conditions. Game dwindled. The crops that had been easily gathered became more difficult to find. The proto-Neolithic Natufian cultures of the Fertile Crescent collapsed. Habitat shrank for man, animal and plants. North Africa became barren. African populations -- most likely those who were ancestral to the Khoi and San peoples -- moved either south toward Central Africa or north into the Levant, the sole way out.

A general depopulation and competition must have taken place, as resources became scarce. In Europe especially there must have been a severe depopulation, with the remaining people concentrating again near the coasts and in mountain valleys, where the climate was somewhat more temperate. This certainly spelled the end of the High Paleolithic Magdalenian cultures. This might have been the period when the Basque peoples -- whom Cavalli-Sforza suggests may be genetically the descendants of the Magdalenian hunters -- found refuge in the Pyrenees.

Perhaps some few tribes began, under climatic pressure, to learn how to domesticate both the animals and plants they had previously hunted and gathered by chance. This was a slow gradual learning process for those who survived.

"The Younger Dryas climatic episode could perhaps have been the external stress, reducing wild food sources, that caused hunter-gatherer societies to turn to cultivation in efforts to increase their after-harvest reserves of grain." (Ofer Bar-Yosef, from Smith, B., The Emergence of Agriculture, 1998)

A diversion into the causes of sudden climate changes:

Excerpted from J. Adams: Sudden (decade-timescale) transitions and short-lived cold and warm phases in the global climate record at:
http://www.esd.ornl.gov/projects/qen/adams1.html

"The circulation of the North Atlantic Ocean is presently seen as playing a major role in either triggering or amplifying rapid climate changes in the historical and recent geological record. The North Atlantic has a peculiar circulation pattern; the north-east trending Gulf Stream carries warm and relatively salty surface water from the Gulf of Mexico up to the seas between Greenland, Iceland and Norway. Upon reaching there, the surface water cools off and (with the combination of being cooler and relatively salty) becomes dense enough to sink into the deep ocean. The 'pull' exerted by this dense sinking water is thought to help maintain the strength of the warm Gulf Stream, ensuring a current of warm tropical water into the North Atlantic that sends mild air masses across to the European continent.

"If the sinking process in the North Atlantic were to diminish or cease, the weakening of the warm Gulf Stream would mean that Europe had colder winters... In the North Atlantic itself, sea ice would form more readily in the cooler winter waters due to a shut-off of the Gulf Stream, and for a greater part of the year the ice would form a continuous lid over the North Atlantic. A lid of sea ice over the North Atlantic would last for a greater proportion of the year; this would reflect back solar heat, leading to cooler summers on the adjacent landmass as well as colder winters. With cooler summers, snow cover would last longer into the spring, further cooling the climate by reflecting back the sun's heat. The rapid result of all this would be a European and west Siberian climate that was substantially colder (because the warm Gulf Stream air was diverted away by the shutting down of the North Atlantic circulation, and by a high-pressure region formed over the sea ice lid) and substantially drier (because the air that reached Europe would carry less moisture, having come from a cold sea ice surface rather than the warm Gulf Stream)."

This rapid-change process that Adams describes could be triggered by a large influx of fresh water into the northern seas, diluting the salty sea water enough to prevent its sinking and flowing southward, thus shutting down the warming Gulf Stream. This could occur either through a period of wet years or as a result of the general climatic warming itself, through the melting of the northern ice. If vast amounts of meltwater were trapped behind ice-dams and then suddenly released, it could rapidly trigger an event such as the Younger Dryas. (In fact, many climatologists believe that the current global warming could trigger such a "switching off" of the North Atlantic thermo-haline conveyor, ironically causing near-glacial conditions in Europe, despite the general warming of the earth's climate.) Correspondingly, the sudden change could also occur in the opposite direction, as a result of increased evaporation in the tropical Atlantic region.

See also:

Steig, E. G., et al: "Synchronous Climate Changes in Antarctica and the North Atlantic", Science, Volume 282, Number 5386, Issue of 2 Oct 1998, pp. 92 - 95

Stocker, Thomas F. "The Seesaw Effect" Science Volume 282, Number 5386, Issue of 2 Oct 1998, pp. 61 - 62

Taylor, K. C. et al, "The Holocene-Younger Dryas Transition Recorded at Summit, Greenland", Science, Volume 278, Number 5339, Issue of 31 October 1997, pp. 825 - 827

Taylor, K., "Rapid Climate Change", American Scientist, Vol. 87, Number 4, July-August 1999; URL@ http://www.sigmaxi.org/amsci/articles/99articles/taylor.html

---

The Holocene

About 11,400 years ago, there was a sudden shift back to warmer conditions. The Younger Dryas ended even more suddenly than it began. It is estimated that, in a period of 20 years, temperatures rose in some regions as much as 5 - 10 degrees Celsius (9 - 18 degrees F). This is a dramatic change that must have had broad implications for the dispersal of human populations and languages. See discussion below, "Did Indo-European Languages spread before farming?" by Adams and Otte)

From 11,400 ya to 8,200 ya climate continued to improve. Warmer, moister conditions returned. Forests expanded in Europe. Deserts decreased to a lesser extent than at present. This would seem to have been a time of opportunity for great expansion by those populations that had survived the Younger Dryas period of stress, whether by developing new techniques or by retreating to refuge areas. This period (the beginning of the

Mesolithic) could well have been the time for the expansion (and fragmentation) of Nostratic out of the Middle East or Pontic areas.

From 9,000 ya to 4,500 ya was the period known as the Holocene Optimum, a period of warmer and moister climate than today. The Saharan and Arabian deserts almost disappeared, becoming grassland. This is the time that saw the rise of the Neolithic and the beginnings of Civilization, of urban, stratified societies. The branching of the linguistic super-families into the current lower order families would date to this period.

However, at 8,200 years ago there occurred a brief (400 year) return to conditions of stress and colder, more arid conditions. This cooling event was about half the severity of the Younger Dryas, and also much briefer. Cold and arid conditions especially affected Northern Europe, Northern Africa and Southern Asia. It is known that many Neolithic settlements in the Middle East were abandoned at this time. This climate change happened very quickly (in less than a generation) at both ends. There would have been little time to adapt at the beginning, and at the end the race for repopulation would have gone to the swift and the opportunistic.

As a example of how in-depth understanding of climate change is accelerating, an article appeared in the July 22, 1999 issue of Nature suggesting that a massive influx of Northern freshwater caused this brief cold event, based on analysis of Greenland icecores. "...The scientists propose that two glacial lakes dammed behind a remnant of the Laurentide Ice Sheet burst through when the ice melted. The freshwater acted like a blanket over the denser salt water of the Labrador Sea, altering circulation and preventing ocean heat from escaping into the atmosphere."

Beginning by 4,500 years ago, conditions began to approximate those of the present, deteriorating somewhat from the Holocene Optimum. The desert again began to encroach upon the savannah, a process that continues to this day. At various times, there still occurred climatic jumps that affected human history (and by this time Prehistory has become History). Examples are a sudden arid event 4,000 years ago covering North Africa and Southern Asia and the Little Ice Age of Medieval times (1400-1800).

Another example of recent research:
(From: Sincell, M., "A Wobbly Start for the Sahara", Science, Volume 285, Number 5426, 16 Jul 1999, p 325)

"Studies of fossilized pollen have shown that grasses and shrubbery covered what is now the Sahara until some unknown environmental catastrophe dried up all the water, leaving nothing but sand. The exact timing is uncertain, but one interpretation of the pollen data suggests that a relatively mild arid episode between 6000 and 7000 years ago was followed by a severe 400-year drought starting 4000 years ago. Such a disaster might have driven entire civilizations out of the desert, leading them to found new societies on the banks of the Nile, the Tigris, and the Euphrates rivers. But the cause of the postulated droughts remained a mystery.

"Climatologist Martin Claussen and co-workers at the Potsdam Institute for Climate Impact Research in Germany are proposing that Earth's changing tilt triggered the rapid drying of the Sahara. Like a spinning top slowly wobbling on its tip, Earth's tilt has decreased from 24.14 degrees to 23.45 degrees in the last 9000 years, resulting in cooler summers in the Northern Hemisphere. When Claussen introduced cooler Northern summers into a computer simulation of Earth's atmosphere, oceans, and vegetation, the monsoon storms that provide water to the Sahara grew weaker, killing off some of the native plants. The initial reduction in vegetation further reduced rainfall, says Claussen, starting a vicious cycle of desertification that began to accelerate about 4000 years ago. Less than 400 years later, Claussen's team found, the drought caused by the vegetation-feedback mechanism could have wiped out almost all plant life in the desert."

.

CLIMATIC IMPLICATIONS FOR LANGUAGE DISPERSAL: Case Studies

The Younger Dryas:

One recent interesting example of a collaborative effort between climate studies and paleolinguistics is an article in Current Anthropology by Jonathan Adams and Marcel Otte, postulating the Younger Dryas cold event as the impetus for the dispersal of the Indo-European language complex.

Adams, J. and Otte, M., "Did Indo-European Languages spread before farming?", Current Anthropology, January 1999. URL@ http://www.esd.ornl.gov/projects/qen/Indo2.html

Abstract:
"The late Glacial record of vegetation and climate suggests that major changes in hunter-gatherer population density might have occurred across Europe and Asia as a result of extreme climate fluctuations. We hypothesize that a reduction in population density across most of the region during the coldest part of the Younger Dryas (around 12,800-11,400 cal. y.a.) may have been followed by a sudden rebound phase, when climate switched back to warm, moist Holocene conditions over only a few decades. A 'sparse wave' of hunter-gatherers migrating rapidly out of a refugial area (possibly located in southern Europe and/or the Near East) would have made a disproportionate contribution to the genetic and linguistic legacy of the region. This may explain part of the initial prehistoric dispersal pattern of the Indo-European languages. Other smaller and somewhat later climate changes, such as the cold event at 8,200 cal. y.a., are also candidates for this process of regional depopulation followed by repopulation from a restricted source region. The possibility should be considered in addition to hypotheses invoking spread of these languages by early farmers or warlike cultures."

Adams and Otte argue that severe and sudden climatic changes forced populations into a few refuges, whether they were in lands further south or in coastal valleys. Such isolation and dispersion would have fostered linguistic drift and change. As conditions improved, those groups that were quickest to adapt to the change would have expanded rapidly through the formerly de-populated regions, carrying their languages with them. They offer Basque as an example of a population that did not respond and expand to the amelioration, leaving them as an isolate in their refugial area. There must have been many other such cases of isolated languages and peoples who withered or were absorbed, such as Etruscan, Ligurian and Iberian, all of which occupied areas that would have been refuges from climatic stress.

"...The paleoenvironmental record suggests various times over the last 15,000 years at which major changes in hunter-gatherer population density could have occurred on a regional scale, due to environmental changes. Such population shifts would be difficult to detect in a sparse archaeological record subject to large C14 anomalies, but they remain a distinct possibility given the magnitude of the climate and ecological changes recorded from across the region. While the ending of the Younger Dryas event seems particularly likely to have resulted in population waves spreading across the region within the approximate time range of the origin of Indo-European languages, any one of these prehistoric changes could have initiated the spread of the Indo-European language group (and in a broader sense the linked Indo-European/Finno-Ugric group). Given the existing dating and the detailed linguistic analysis which suggests a divergence time around 7,000 cal. y.a., a somewhat later climate change (early-to-mid Holocene; e.g. the 8,200 cal. y.a. or the 5,900 cal. y.a. cold events) would seem to rest more easily with observations. An 8,200 y.a. change could have promoted spread of Indo-European languages by either hunter-gatherers, farmers, or both. "

The Black Sea Flood:

Another instance where climatic "catastrophism" has been drawn upon to explain dramatic human population and linguistic dispersal is the subject of a recent book Noah's Flood: The New Scientific Discoveries About the Event That Changed History, by William Ryan and Walter Pitman, geoscientists at Lamont-Doherty Laboratories. The sensationalism of the title belies the serious research that underlies the thesis of the book, which is that in the matter of a few years -- 7600 years ago -- the Black Sea changed from a freshwater lake, with a rich ecology, to a barren salt water sea by means of a sudden flood that raised the water level by 300 feet in a few years.

Original report of research (1997):
"Based on analyses of Black Sea sediments, oceanographers William Ryan and Walter Pitman of the Lamont-Doherty Earth Observatory in Palisades, New York, have put together evidence that about 7500 years ago, this great deluge [the prototype of Noah's Flood] really happened, suddenly filling the Black Sea to its present level....

"...Ryan and Pitman [propose] that the flood also fostered the spread of agriculture across Neolithic Europe. By 9000 years ago, farming--both cultivating grains and raising livestock--had originated in southwestern Asia; by 8000 years ago, it had spread to Greece and into the Balkans, including Romania and Bulgaria. Farming stayed in this region for some centuries, then surged across eastern Europe and into central Europe east of the Rhine River at about the same time as the flood, Bailey notes. Archaeologists debate whether the migration of people or the passing of seeds and animals from neighbor to neighbor drove the dispersion of farming. Pitman and Ryan argue for mass migration.

"We would say this flood caused a diaspora," says Pitman. The timing is right, he says, to have driven Neolithic farmers up the rich river valleys into central Europe, as well as Egypt and southern Mesopotamia, where a new and distinctive farming culture appears at about that time. In the Mesopotamian kingdoms, the shaken immigrants' tales might have grown into the Sumerian flood myth and eventually evolved into the biblical flood, he suggests."
(Kerr, R. Science, V. 279, N. 5354, pg. 1132)

During the Ice Ages, at a time of great aridity, the Black Sea was a small freshwater lake cut off from the lowered ocean levels. As the glaciers began to melt, ca. 14,500 years ago, the meltwater poured into the Central Asian "Great Lakes", then into the Black Sea, and breached through at the Sea of Marmara and on to the Aegean and the Mediterranean. As a result, the freshwater Black Sea grew to twice its current extent. This first meltdown ended by 12,000 years ago with the Younger Dryas, and the connection between the Black and Mediterranean Seas became dammed. The flow into the Black Sea diminished, and it slowly began to shrink. The great Central Asian rivers dried up, and the Black Sea must have become a refuge and an oasis of fresh water and varied flora and fauna resources throughout the dry years of the YD.

The second glacial melting begins after the Younger Dryas, about 11,400 ya. The ocean levels again began to rise. However, the meltwater that had flowed south toward the Black Sea -- blocked to the North by the glacial mass -- now flowed into the North and Arctic Seas as the glaciers receded. The Black Sea continued to shrink, until by 8600 ya it was some 350 feet below the level of the ocean (and the Mediterranean), and it was separated from the Aegean arm of the Mediterranean by a dam at the Bosphorus. But, smaller though it had become, it would still have been an oasis in a dry land, fertile and heavily populated, especially during the cold and arid climatic event of 8200 yrs ago.

Then the dam broke, as the ocean level finally rose above it. The floodwaters rapidly poured in as a massive waterfall, daily widening the breach. All that had supported human life fled or succumbed to the salt. Coastal villages drowned in a matter of weeks as the sea level rose.

This process is not disputed. Ryan and Pitman speculate the Black Sea area had been the mixing ground of the proto-languages of Indo-European, Kartvelian, Semitic, and Sumerian, and that when the dam broke and the region flooded, there was an out-migration that established the new cultures in Eastern Europe (Vinca and proto-IE), Anatolia, the Levant, Sumeria and Egypt.

For Ryan and Pitman, who are geoscientists not historical linguists, this flood seems to resolve in one great event most of the enduring linguistic issues of all Western Eurasia. There is certainly much to dispute in these speculations. The Neolithic was well underway in Mesopotamia before the flood, so it is hard to imagine that this event was ancestral to the rise of the Sumerians. Similarly, getting Semitic peoples (or "proto-Egyptians") to the Black Sea at that date seems to contradict what is known about the development and dispersal of Afro-Asiatic. Ryan and Pitman base much of their speculation on the pervasiveness of the Myth of the Great Flood, which they consider to be a memory of the Black Sea Flood.

But the fact of the event itself of the Black Sea Flood is not in question, nor can there be any doubt that it must have had vast influence on language dispersal over a large area. I would imagine that any student of the dispersal of Nostratic/Eurasiatic superphyla, or of IE or of the Caucasian families -- to use a few examples -- must take this event into account. The flood and the movement of peoples out of the impacted region must have had important linguistic effects in a large arc from Anatolia and SE Europe through the Caucasus to the Caspian, at the least.

See also:
Kerr, R., "Black Sea Deluge May Have Helped Spread Farming", Science", Volume 279, Number 5354, Issue of 20 Feb. 1998, p 1132

Ryan, W. and Pitman, W. "Letter - Response" Science, Volume 280, Number 5363, Issue of 24 Apr. 1998, p 499

Some Comments on Nostratic

Following is an Email discussion I had with Alan Bomhard two years ago when I was beginning to consider the climate/language picture. Alan was willing to attach dates to these early language dispersals. It is to be noted that the time-frame he regards as critical to Nostratic and Afro-Asiatic is precisely the period of the greatest climatic fluctuations: from the end of the last glacial (14,000 ya - 12,800 ya) through the Younger Dryas (12,800 ya - 11,400 ya) to the beginning of the Holocene warming (11,400 - 8,200), which was interrupted at 8,200 ya by a period of climatic reversion to semi-glacial conditions.

Communication from Alan Bomhard: *[My notes appear in italics… RF]*
"In my opinion, [John C.] Kerns has hit the nail on the head (Bomhard-Kerns 1994:155): 'I believe that the Mesolithic culture, with its Nostratic language, had its beginning in or near

the Fertile Crescent just south of the Caucasus'. Let us now reexamine the evidence from the Nostratic daughter languages and see how it leads to this conclusion.

"The Indo-European homeland was most likely to the north of and between the Black and Caspian Seas. However, Joanna Nichols has convincingly argued that Pre-Indo-European originated in Central Asia and later spread westward to the North Pontic/Steppe zone that was the geographical location where Proto-Indo-European proper developed, where it began to split up into different dialect groups, and from which its descendants spread into Europe, the Iranian plateau, and northern India. Likewise, again as argued by Nichols, Pre-Uralic may be presumed to have originated in Central Asia and to have spread westward, following a more northerly route than Pre-Indo-European. Thus, it is likely that the Eurasiatic parent language was located in Central Asia and that it is to be dated roughly at about 9,000 BCE. This would mean that the eastern Eurasiatic languages (Altaic, Chukchi-Kamchatkan, Gilyak, and Eskimo-Aleut) must have spread eastward from Central Asia (more specifically, the area traditionally called "Western Turkestan") to their prehistoric homelands. Nichols has also speculated that Pre-Kartvelian may have originally been located in Central Asia, from which it spread westward along a southern route below the Caspian Sea to the Caucasus Mountains. The Elamo-Dravidian homeland may be placed roughly in western and central modern-day Iran and dated at about 8,000 BCE. Finally, the homeland of Afroasiatic may be placed in the Middle East in the Levant and dated at about 10,000 BCE. Working backwards geographically and chronologically, we arrive at the only possible homeland for Proto-Nostratic, namely, "the Fertile Crescent just south of the Caucasus".

"Thus, the following scenario emerges: the unified Nostratic parent language may be dated to between 15,000 to 12,000 BCE, that is, at the end of the last Ice Age -- it was located in the Fertile Crescent just south of the Caucasus.
RF: During the Ice Age, the northern Fertile Crescent would have been a refuge from the pervasive cold and aridity. The upland valleys were better watered, more sheltered and less frigid than were the lowlands. These early Nostratic folk may not have been great in numbers, considering the population constrictions of the Glacial Maximum, and they were likely divided into isolated valley bands.

"Beginning around 12,000 BCE, Nostratic began to expand, and, by 10,000 BCE, several distinct dialect groups had appeared. "
RF: Climate improved radically and quickly ca. 14,000 ya. (12,000 BCE) The Fertile Crescent and Northern Africa developed into well-watered savannah, optimal for hunter/gather cultures. Population expanded into these areas because of the increased food supply. Population also expanded into the unpopulated northern regions. The Mesolithic, post-glacial culture spread throughout southern Eurasia. As Europe became woodland, the Magdalenian hunters began to follow the herds through Northern Eurasia. Central Asia was rich and well-watered as a result of the glacial meltwater.

"The first to split off was Afroasiatic. One dialect group spread from the Fertile Crescent to the northeast, eventually reaching Central Asia some time before 9,000 BCE -- this was

Eurasiatic. Another dialect group spread eastward into western and central Iran, where it developed into Elamo-Dravidian at about 8,000 BCE. If Nichols is correct in seeing Pre-Kartvelian as having migrated from Central Asia westward below the Caspian Sea to the Caucasus, this would seem to imply that Pre-Kartvelian had first migrated northeastward from the Fertile Crescent along with or as part of Pre-Eurasiatic, that it stopped somewhere along the way, and that it then returned to the Middle East."
RF: Expansion came with the warming, and then bottleneck re-appeared with the Younger Dryas, at 10,800 BCE. Near-glacial conditions of cold and aridity returned. This was a time of great stress to the previously expanded populations, a time of out-migration and depopulation – and of linguistic dispersal. Habitat shrank for man, animal and grass. North Africa became barren. African populations moved either south to central Africa, or north into the Levant and Fertile Crescent. Many language families also must have become extinct during this time of stress, and many linguistic links broken. Expansion resumed again about 9,800 BCE, coupled with the beginnings of domestication.

9,000 - 6,200 BCE: There were northward and eastward movements of Eurasiatic, Elamo-Dravidian and Kartvelian (Bomhard). Climate improved again, during the beginning of the Holocene Optimum, a climate more beneficent than today's. Expansion continued into the areas that had been depopulated, particularly into Europe and Central Asia. Populations in situ also expanded. One piece of archaeological evidence for this is the rise of true agricultural cultures in the Fertile Crescent, involving the domestication of wheat and barley, as well as the husbandry of sheep, goats and swine. Whether domestication began as a response to Younger Dryas stress -- as Bar-Yosef suggests -- or was unleashed by its termination, this period was certainly the time of the rise of the agricultural Neolithic. Agriculture allowed for population expansion, which of itself creates out-migration -- not necessarily invasion, but rather a natural expansion of peoples to new and open land. The peoples who were once "Nostratic", who had learned the domestication of wheat and barley, of pigs and goats, expanded in all directions.

6,200 - 6,000 BCE: A new period of climatic stress came with this semi-glacial event circa 8,200 years ago, which might have been devastating to peoples who had adapted to the warmer climate. When that event ended a few hundred years later, there may have been an opening for another movement from the south into a sparsely inhabited Europe, as well as to areas where aridity had devastated the nascent agriculture.

6,000 BCE: Climate returned to the warm state. Expansion resumed, in all directions. The Black Sea Flood caused flight over a wide area. The Nostratic descendants, who had been successful in evolving out of a strictly opportunistic culture, continued to move into the vacuum of Southern Eurasia, in all directions. Population increased steadily. The Fertile Crescent was then in a state of optimal climate. In Sumeria, and later in the Nile and Indus valleys, population pressure and the success of agriculture brought about the rise of the State, the beginning of Civilization, and the beginning of History.

Conclusion:

The notes above constitute but a few suggestions as to how climatic studies can be used -- and must be taken into account -- in any discussion of language dispersal. I attempt to draw no conclusions at this point other than that climate was a constantly and radically changing variable that determined the ecology and resources upon which man depended throughout the Paleolithic and Neolithic. Such conclusions are for further papers, more narrowly focussed. I have barely touched on areas outside the range of Nostratic and Eurasiatic, such as South and East Asia, as well as the Americas, areas where I also believe that climatic correlations can be of great use to the understanding of linguistic dispersal.

Climate was a driving force that determined where and how humans could survive, and it greatly influenced how they and their languages would move, disperse or disappear. This is especially true everywhere north of the Tropics, the primary areas of human expansion over the past 60,000 years. Neil Boaz has used the phrase "the climate pump" to describe how climatic fluctuations moved Homo Erectus and Early Homo Sapiens in and out of Africa, but that influence did not end when humans began to use clothing and fire and boats -- or agriculture. Climate was not a background detail, but was more often the prime impetus behind population movements, often changing more quickly than humans could adapt.

It is worth emphasizing here that much of what is known about the details of climate change is the result of recent research. Ryan and Pitman put together the details of the Black Sea flood less than two years ago, as a result of deep-water sonar explorations undertaken over the past decade. The details of the Younger Dryas period -- particularly the speed at which climate changed -- are a product of on-going studies of the Greenland ice cores, supplemented by other research throughout the world, including palynology (pollen studies) and dendro-chronology (tree-ring dating). The understanding of sudden climate change is very recent, much of it the result of concern about the "greenhouse effect", which may not be as gradual (or even beneficent) as people had thought. (For this, refer to the work of Wallace Broeker or Kendrick Taylor)

Just as current DNA research is constantly adding to -- and altering -- the understanding of population movements that are at the heart of language change, so also is climate study. It is the task of paleo-linguists to use this new understanding. This is the essence of what Lord Renfrew called "The Emerging Synthesis", a cornerstone of ASLIP, and I believe that paleo-climatic research can be as critical to this synthesis as are genetics and archaeology.

Preface to "Bipeds, Tools, and Speech"

By Roger W. Wescott, *Mother Tongue* Editor

Since 1929, there has been a minoritarian view of human evolution known as *aquaticism,* which denies that our ancestors moved directly from Miocene forests to Pleistocene grasslands. Instead, it pictures them as having detoured through shallow Pliocene waters, till they became furless and bipedal.

Most evidence of this transition is regarded by aquaticists as having been water-buried. The extraordinary brain growth of hominines – comparable to that of cetaceans – they explain as due largely to a rich diet of fish, molluscs, and crustaceans.

And our vocal versatility, culminating in speech, they see as compensation for the declining reliability of visual and olfactory signaling in an aqueous environment. In this article, Verhaegen and Munro present a fine overview of recent data supporting the link between hominine aquaticism and language.

* * *

Bipeds, Tools and Speech

By Marc Verhaegen & Stephen Munro

Studiecentrum Antropologie
Mechelbaan 338, 2580 Putte, Belgium
6 October 1999

After a lecture at the symposium "Water and human evolution", Ghent, Belgium, 29-30 April 1999
http://allserv.rug.ac.be/~mvaneech/Programme.html

There is general consensus, within certain sections of the anthropological community, that australopithecines were ancestral to humans, meaning they had already separated from the ancestral line leading to chimpanzees. Evidence suggesting that australopithecines were bipedal, such as fossilized footprints and skeletal remains, is often used as supporting evidence for this hypothesis.

Of course, this hypothesis is based, at least partly, on the assumption that chimpanzees and humans descended from ancestors that were not yet bipedal, and that bipedalism emerged only after the ancestral lines leading to chimpanzees and humans had separated. The once popular "savanna theory" held that hominid bipedalism emerged when human ancestors moved from the forests into a more open and arid environment.

Here, we outline an alternative hypothesis, which we believe more accurately reflects the available evidence. We believe, contrary to the popularly held view described above, that the australopithecines were no more closely related to the ancestors of humans than they were to the ancestors of the African apes. Further, we propose that the common ancestor of humans, chimps and gorillas was already at least partly bipedal, regularly wading, possibly in coastal mangrove forests, in much the same way as proboscis monkeys do today. Gorillas and chimpanzees, according to this hypothesis, evolved knuckle-walking features independently, in parallel, after moving from the coast to the African interior by following rivers and gallery forests, or by migrating with expanding forests.

Humans, according to this theory, descended from a hominid population that remained nearer the coast, perhaps in flooded mangrove or near-coast forests where they waded regularly. Eventually, they became better divers too, giving rise to the big-brained, long-returned to a more terrestrial lifestyle, becoming, essentially, long-legged terrestrial bipeds who retained very good swimming and diving adaptations. This hypothesis, in our opinion, helps explain many unique human adaptations, including the development of human tool manufacturing skills, and the origins of speech.

Hominid fossils and scenarios

The evolutionary history of all animals, including our ancestors, has been influenced by a number of environmental factors. Thus, we believe many evolutionary insights can be gained by comparing the parallel and convergent adaptations of different animals living in similar environments. In fact, we believe evidence gained from the comparative anatomy and physiology of living species is as important to evolutionary studies as the fossil evidence is.

The fossil record displays well-known shortcomings. It is biased and incomplete. For example, it is possible that hominids living in certain environments were less likely to leave fossilized remains than hominids living in other milieus. More specifically, geologists note that fossilization is extremely difficult in mangrove areas, where tidal water movements can spread the bones over a vast area, and where the high acidity of the water might increase the likelihood of the bones dissolving. Moreover, in mangrove areas the sea floor is flat, so there is almost no chance that a landslide would ever cover remains.

Because of the scantiness of the fossil record, paleontologists now generally accept the late Colin Patterson's view that the direct ancestors of living species are unlikely to ever be found within it (Nelson, 1998). As a result, it is probable that most, if not all, fossil hominid species found to date are simply extinct side-branches of the lines leading to the present living hominids. In part, it was this likelihood that led us to be cautious about using the fossil record as the sole basis for attempting to develop a viable hominid "family tree". Instead, we considered all the available evidence, including that gained from comparative anatomy and physiology.

Wading in forest swamps?

Most primates are four-legged tree-dwellers with a tendency to truncal erectness and with very mobile limb joints which enable them to reach, climb and leap through trees. Because of this locomotor flexibility, they can, more easily than typical quadruped land mammals, adopt a bipedal gait when necessary.

Many primates adopt a bipedal gait when wading through water. For example, lowland gorillas have been observed wading on their hind limbs through forest swamps in search of edible sedges and aquatic herbs (Chadwik, 1995; Doran and McNeilage, 1997). The mangrove-dwelling proboscis monkeys also cross stretches of water to move from one mangrove tree to another, and walk on two legs when making these treks. In fact, they have even been seen using bipedal locomotion on dry ground (Morgan, 1997).

Note that this bipedal wading gait is different from the hopping bipedalism used by other primates and mammals when moving on the ground (for instance, tarsiers, indris, jerboas and kangaroos). This latter gait incorporates bent knees and hips rather than the more linear stature preferred for wading. One advantage of the erect wading posture is that it allows primates to hold their body, arms and head as far as possible above the water surface, allowing them to use their hands to search for and manipulate food. Bears have been seen searching for food in water while wading bipedally.

Rightly or wrongly, most anthropologists still base their estimates of when human bipedalism emerged almost exclusively on the available fossil evidence. Up until a few years ago, this evidence was used to suggest that bipedalism arose some four million years ago, in a savanna environment. Recent studies, however, as well as fossil finds such as *Ardipithecus ramidus* and *Australopithecus anamensis*, have forced a reconsideration of this traditional view. Today, most anthropologists accept the view that bipedalism probably emerged earlier, in a wooded or forested habitat (Tobias, 1998).

We believe the common ancestors of chimpanzees, gorillas and humans frequently waded and climbed in seasonally or permanently flooded forests, perhaps in mangrove forests on the coasts of the ancient Tethys Sea, somewhere between Afro-Arabia and Eurasia, between what is now the Mediterranean and the Arabian Sea. There is presently no evidence to discount the idea that hominid bipedalism evolved in a milieu where both trees and water were present. Most, if not all, early hominid fossils have been discovered in areas which were then well forested and close to water (Radosevich *et al.*, 1992; Rayner, 1993). It is important to note, however, that although in our opinion the last common hominid ancestor was a regular wader, it almost certainly continued to use the trees for refuge, sleep and food gathering.

Emergence of australopiths and African apes

Humans belong to the hominoid biological group, which consists of the lesser apes (hylobatids) and the great apes (pongids and hominids). Today, hylobatids (gibbons and siamangs) and pongids (orangutans) live in Asia, whereas hominids (humans, chimpanzees and gorillas) live in Africa. According to the fossil record, however, great apes that had pongid and/or hominid features, such as *Dryopithecus*, *Graecopithecus*, *Ankarapithecus* and *Sivapithecus*, lived in Europe, Anatolia and India between 14 and 8 million years ago (Andrews, 1995; Algaput *et al.*, 1996). This suggests that the ancestors of Asian pongids and African hominids most probably lived somewhere between the fossil-bearing regions of Africa, Europe and India, possibly in what is now the Middle East (Stewart and Disotell, 1998), perhaps near one of the then abundant seas which formed as a result of the desiccation and fragmentation of the Tethys Sea.

In 1974, the partial maxilla and teeth of the 17-million-year-old *Heliopithecus* were collected at Al-Dabtiyah, eastern Saudi Arabia, between continental sandstone-clays and marine limestone-marls: "They appear to be the earliest occurrence of dryopithecines found at any distance from the East African early Miocene sites and their habitat seems to have been near to the tropical shore of the Tethys epi-continental sea" (Whybrow and Bassiouni, 1986). This Saudi ape may have been the first recognizable member of the great-ape branch (Andrews *et al.*, 1987). The oldest European dryopithecine-like fossils, *Austriacopithecus* from Dèvinska Nová Ves and Kleinhadersdorf near the Slovakian-Austrian border, about 14 million years old, also lay in marine nearshore sands, and *Droyopithecus* fossils 13-10 million years old derive from what were then swampy forests (Steininger, 1986; Mein, 1986).

We believe a basic great ape population may have clustered somewhere around the ancient Tethys Sea, which once extended from the Mediterranean to Northern India. This cluster may have given rise to different offshoots that entered Europe, Asia and Africa by following rivers and gallery forests upstream, or by migrating with expanding forests. These migrations, in our opinion, led to the fossil and living great apes; in Eurasia, the dryopithecines and the orang-utan; in Africa, the australopithecines and the African apes. Meanwhile, part of the population remained in forests near the coast, where they eventually became efficient swimmers and divers. Some of these species later returned to the land and became predominantly terrestrial bipeds.

According to molecular evidence, the great apes split into pongids and hominids some 14 to 10 million years ago. The ancestral line leading to the gorillas separated from the line leading to humans and chimpanzees about 8 to 6 million years ago, and the ancestors of chimpanzees and humans separated between about 6 and 4 million years ago.

We believe the ancestral line leading to the gorilla branched off from the stem hominid population when it moved from near coastal forests into the African interior, perhaps by

Figures

Figure 1: Miocene great ape fossils:
Africa ca.20-14 mya: *Proconsul*, *Morotopithecus*, *Kenyapithecus*…
Europe-Anatolia ca.12-8 mya: *Dryo-*, *Graeco-*, *Ankarapithecus*…
India ca.12-8 mya: *Sivapithecus*

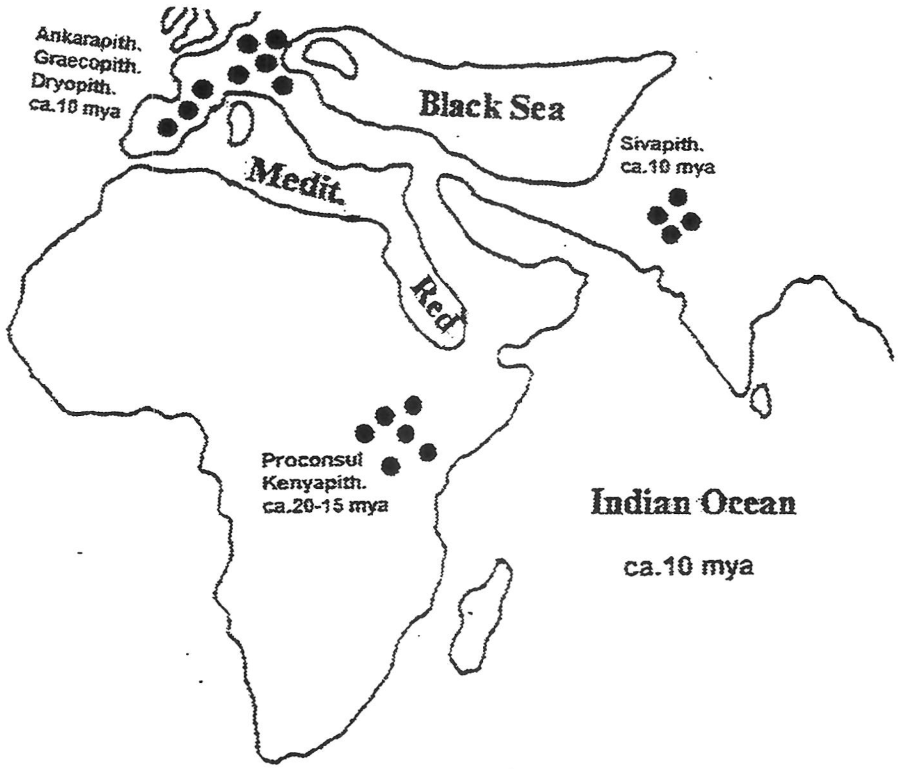
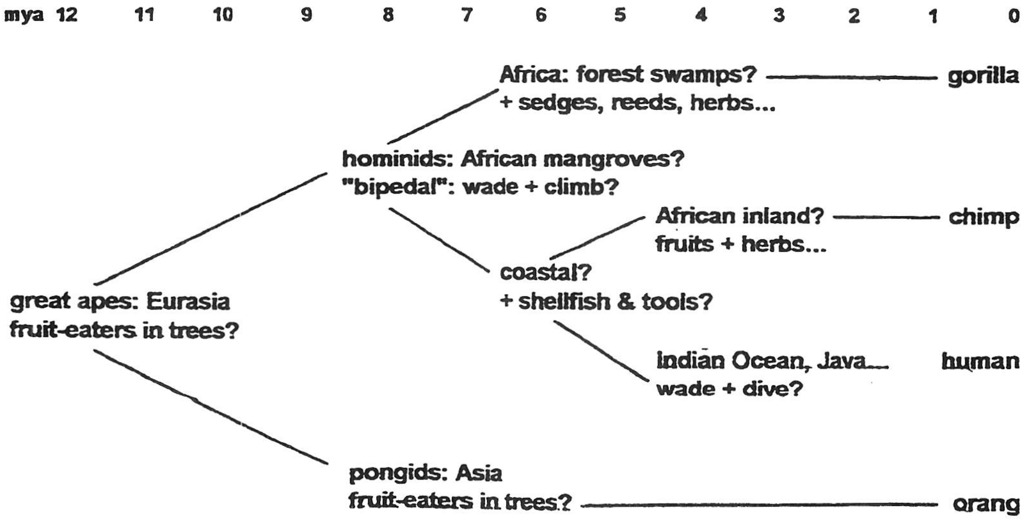

Ankarapith.
Graecopith.
Dryopith.
ca.10 mya
Black Sea
Medit.
Red
Sivapith.
ca.10 mya
Proconsul
Kenyapith.
ca.20-15 mya
Indian Ocean
ca.10 mya

Figure 2: Evolutionary tree of hominids: hypothetical and schematic reconstruction of African ape and human evolution, based on comparative anatomy and behavior, geographical distribution and biomolecular data of living hominids. Fossil species *Ardipithecus*, *Australopithecus* and *Homo* are sidebranches of the lines leading to the living hominids.


mya 12
11
10
9
8
7
6
5
4
3
2
1
0
Africa: forest swamps?
+ sedges, reeds, herbs...
gorilla
hominids: African mangroves?
"bipedal": wade + climb?
African inland?
fruits + herbs...
chimp
coastal?
+ shellfish & tools?
great apes: Eurasia
fruit-eaters in trees?
Indian Ocean, Java...
wade + dive?
human
pongids: Asia
fruit-eaters in trees?
orang

following the rivers and gallery forests of the African Rift Valley. It is not impossible that this ancestral line might have given rise to the very large australopithecine species such as *Australopithecus boisei* (Kleindienst, 1975).

The ancestral chimpanzee population probably migrated from the coast to the African interior a few million years after the gorilla, perhaps by also following rivers and gallery forests. It is likely that the inland hominid branches would have evolved in parallel to some extent. As they migrated further inland, shellfish would have become rarer, and therefore, other plant and animal food would have been required to replace any shellfish in the diet. The inland populations would have become more herbivorous and spent less and less time in the water. The ancestors of chimpanzees, according to this hypothesis, stayed longer at the coast than the ancestors of gorillas, and we would therefore expect chimpanzees to be more omnivorous than gorillas, because the opportunity for harvesting shellfish, for example, would have been greater on the coast than in more inland milieus.

Emergence of the *Homo* genus

The hominid populations that remained in forested areas near the coast would naturally have become more efficient at exploiting the available aquatic resources. Initially, they may have partly fed on the oysters fixed to the mangrove trunks exposed at low tide, and later, they may have developed adaptations which allowed them to dive for shellfish. Note that the long-chain poly-unsaturated lipid ratios of tropical fish and shellfish are more similar to the ratios in the human brain than any other food source known (Broadhurst *et al.*, 1998). We believe these coastal hominid populations may have given rise to species of the *Homo* genus; large-brained, long-legged creatures, able to wade, swim and dive efficiently, and increasingly less reliant on trees.

Presumably, these coastal hominids had rather thick tooth enamel (Martin, 1985). Some earlier (*Graeco-* and *Ankarapithecus*) and later (*Australopithecus*) hominoids actually had superthick enamel. Enamel is extremely hard, and thick enamel is typical of species like orangutans, capuchin monkeys and sea otters that consume hard-shelled fruits, nuts or molluscs. Walker (1981) even wrote: "If, for example, a mammalogist who knows nothing about hominids were asked which mammalian molar most resembled those of *Australopithecus*, the answer would probably be the molars of the sea otter (*Enhydra lutris*). This species possesses small anterior teeth, and large, flat molars with thick enamel." Walker believes the thick enamel in sea otters is probably not for cracking shells, but rather to protect against the occasional hard inclusions inside the shells which would otherwise damage the dentition.

Tool use is seen in many animals, though perhaps the most obvious mammalian examples, with the exception of humans, are capuchin monkeys, chimpanzees and sea otters. They all try to open hard-shelled foods by hammering them with hard objects. Sea otters, for instance, crack open shellfish with stones while floating on their backs. Capuchins crack open nuts and oysters with stones, and Fernandes (1991) has even reported that mangrove capuchins use oyster shells where stones are not available. It seems likely, therefore, that tool use in hominids may have begun with shellfish or nut eating. Chimpanzees crack open nuts with stones, including the hard nuts of oil-palm trees (Sugiyama, 1997). Palm trees generally have branchless stems, and as such are easily climbed by bipeds. They could have provided a refuge from predators such as large cats and wild dogs. Palm trees probably grew along the coasts of the Tethys Sea and the Indian Ocean, and coconuts, which can be easily opened by primates using tools, may have provided a valuable source of fresh drinking water (Hugh Harries, personal communication).

Some early hominids may have used hard objects as tools to remove or crack open the shellfish that grew on the trunks of the mangrove trees in the forests in which they lived. For the hominids that migrated inland, to areas where shellfish were less common, tool use probably continued for activities such as nut cracking and termite fishing. Those nearer the coast would have continued to use tools to crack open shellfish, and perhaps the extraordinary stone tool industries we associate with the various *Homo* species is a reflection of their long association with coastal milieus.

Diving as a prerequisite for speech

Today, breath-hold diving is practiced by several subsistence human cultures that gather shellfish or seaweed. Voluntary breath control, a prerequisite for such activity, is an aquatic or marine adaptation, and this is clearly demonstrated when one compares the human respiratory system with that of non-human primates and aquatic mammals. All frequently diving mammals have the ability to take a deep breath whenever they intend to dive. Many of them, like dolphins and seals, also have larger brains than land mammals of equal size. Diving requires a voluntary control of breathing by the central nervous system. In contrast with land mammals, diving species must be able to hyperventilate whenever they *intend* to dive, and hold their breath underwater, just when their oxygen needs are highest.

Many primates and arboreal animals, like gibbons and birds, have an aptitude for vocal and musical expression. They share this capacity for making and interpreting a wide range of sounds with many marine mammals, such as dolphins and humpback whales. Vocalization may have been an important communicative device in an aquatic environment, where more traditional devices such as smell and body language may have been less effective. Perhaps the musical abilities of our arboreal ancestors, combined with our breath-holding skills, a large brain, and the need to communicate in an aquatic environment, were some of the prerequisites for human speech.

Between aquatic and terrestrial

The available evidence suggests that the *Homo* genus evolved from forest-dwelling species who were also part-time bipedal waders, who gradually became more suited to swimming and diving. These species developed a more streamlined and linear body, longer legs and a larger brain. Eventually, they colonized coastal areas and river valleys in Asia, Africa and Europe, where they probably used their tool-using skills to exploit other available resources such as the scavenged carcasses of hippos and other mammals. It is even possible that they hunted these or other animals.

The Javanese Mojokerto fossil, discovered in a river delta amid marine and freshwater molluscs (Ninkovich and Burckle, 1978), and the Georgian Dmanisi fossil, discovered near a lake or pond rich in lacustrine resources (David Lordkipanidze, personal communication), which have both been dated at about 1.8 million years old, might well be the oldest *Homo erectus* fossils ever discovered. There is even archeological evidence that *Homo erectus* reached the island of Flores, in Southeast Asia, some 800,000 years ago, well before any evidence of boat building appears in the archeological record (Morwood *et al.*, 1998; Tobias, 1998). It is likely that *Homo erectus*'s efficient swimming skills, as well as their probable ability to use naturally floating objects as swimming and floating aides, enabled them to cross the great natural water barrier now known as Wallace's Line.

Eventually, at least some of these long-legged hominid species returned to a more terrestrial existence. But just as no *Homo* species may ever have ceased being at least partly terrestrial, it is possible that no *Homo* species ever ceased being at least partly aquatic either. When members of the *Homo* genus did return to a more terrestrial existence, they were unable to revert to knuckle-walking quadrupedalism, like gorillas and chimpanzees, because, whereas gorillas and chimpanzees evolved from short-legged climbing-and-wading aquarborealists, *Homo* was already a long-legged wader and diver with a more linear build.

Terrestrial bipedalism, as seen in humans, has several disadvantages. It is a slower form of locomotion than quadrupedalism, is more conspicuous, less energy-efficient and leads to many ailments such as backaches, hip and knee problems. However, it also has many advantages, mainly associated with the freeing-up of the hands so that they can be used more effectively for communication, and to carry food, water, babies and tools.

We believe this semi-aquatic phase helps explain many unique human adaptations such as furlessness, subcutaneous fat and voluntary breath control – features unique among primates, but common among various water mammals such as seacows, hippopotamuses, walruses, dolphins and whales. It may also help explain why we are much more efficient swimmers and divers than other primates (Bender, 1999; Schagatay, 1996).

In our opinion, it should not be a question whether members of the *Homo* genus were ever aquatic, but rather to what extent were they aquatic. For example, the fossilized remains of Neandertals, traditionally viewed as fully terrestrial hominids, have been discovered exclusively next to coastlines and rivers. The presence of ear exostoses (bony outgrowths of the ear canal, a condition only seen after life-long diving in modern humans) is evidence that at least some Neandertal individuals practiced frequent diving, and traces of cattails on some stone tools suggest their diet might have included aquatic plants (Shreeve, 1995). We believe Neandertals evolved from even more water-based *Homo erectus* populations that moved up the rivers from the coasts into the Eurasian interior. Some Neandertal populations may have retained or re-evolved elements of a wading and diving existence, like some modern human populations such as the Korean Ama.

Conclusion

Our hypothesis is that the last common ancestors of the African hominids lived possibly in coastal mangrove forests, where they waded bipedally and were omnivorous, supplementing their mainly herbi-frugivorous diet with shellfish and other marine resources. This stem population, in our opinion, gave rise to populations which remained near the coast, and to other populations which migrated inland by following rivers or expanding forests. These inland populations probably included the ancestors of the African apes and the various australopithecine species.

The populations that remained near the coast, due to geological factors, probably left fewer fossilized remains. They gave rise to the various species of the *Homo* genus - long-legged, big-brained hominids, very capable swimmers and divers, able to take full advantage of the available resources associated with a coastal milieu. These hominids populated suitable coastal regions and followed rivers into the interiors of Asia, Africa and Europe.

This hypothesis offers a possible scenario as to where and why our ancestors evolved the way they did. It is detailed enough that it can be tested against new evidence as it becomes available, and, in our opinion, can also be used as a predictive tool. As such, its success or failure should either confirm or negate its value as a possible framework for human evolutionary discussion.

References

Algaput, B., Andrews, P., Fortelius, M., Kappelman, J., Temizsoy I., Celebi, H. & Lindsay, W. (1996) A new specimen of *Ankarapithecus meteai* from the Sinap Formation of central Anatolia. *Nature*, 382: 349-351.

Andrews, P. (1995) Ecological apes and ancestors. *Nature*, 376: 555-556.

Andrews P., Martin L. & Whybrow P (1987) Earliest known member of the great ape and human clade. *American Journal of Physical Anthropology*, 72:174-175.

Bender, R. (1999) *Die evolutionsbiologische Grundlage des menschlichen Schwimmens, Tauchens und Watens*. Bern, Switzerland: University of Bern.

Broadhurst, C. L., Cunnane, S. C. & Crawford, M. A. (1998) Rift Valley fish and shellfish provided brain-specific nutrition for early Homo. *British Journal of Nutrition*, 79: 3-21.

Chadwik, D. (1995) Ndoki – last place on earth. *National Geographic*, 188: 2-43.

Doran, D. M. & McNeilage, A. (1997) Gorilla ecology and behavior. *Evolutionary Anthropology*, 6: 120-130.

Fernandes, M. E. B. (1991) Tool use and predation of oysters by the tufted capuchin in brackish water mangrove swamp. *Primates*, 32: 529-531.

Kleindienst, M. R. (1975) On new perspectives on ape and human evolution. *Current Anthropology*, 16: 644-646.

Martin, L. B. (1985) Significance of enamel thickness in hominoid evolution. *Nature*, 314: 260-263.

Mein, P. (1986) Chronological succession of hominoids in the European Neogene. In Else, J. G. & Lee, P. C., eds. *Primate Evolution*, pp. 59-70. Cambridge: Cambridge University Press.

Morgan, E. (1997) *The aquatic ape hypothesis: the most credible theory of human evolution*. London: Souvenir.

Morwood, M. J., O'Sullivan, P. B., Aziz, F. & Raza, A. (1998) Fission-track ages of stone tools and fossils on the east Indonesian island of Flores. *Nature*, 392: 173-6.

Nelson, G. (1998) Colin Paterson (1933-1998). *Nature*, 394: 626.

Radosevich, S. C., Retallack, G. J. & Taieb, M. (1992) Reassessment of the paleoenvironment and preservation of hominid fossils from Hadar, Ethiopia. *American Journal of Physical Anthropology*, 87: 15-27.

Rayner, R. J., Moon, B. P. & Masters, J. C. (1993) The Makapansgat australopithecine environment. *Journal of Human Evolution*, 24: 219-231.

Schagatay, E. (1996) *The human diving response: effects of temperature and training*. Lund, Sweden: University of Lund.

Shreeve, J. (1995) *The Neandertal Enigma: solving the mystery of modern human origins*. New York: Morrow.

Steininger, F. F. (1986) Dating the Paratethys Miocene hominoid record. In Else, J. G. & Lee, P. C., eds. *Primate Evolution*, pp. 71-84. Cambridge: Cambridge University Press.

Stewart, C.-B. & Disotell, T. R. (1998) Primate evolution: in and out of Africa. *Current Biology*, 8: R582-588.

Sugiyama, Y. (1997) Social tradition and the use of tool-composites by wild chimpanzees. *Evolutionary Anthropology*, 6: 23-27.

Tobias, P. V. (1998) Water and Human Evolution. *Out There*, 35: 38-44.

Walker, A. (1981) Diet and teeth – dietary hypotheses and human evolution. *Philosophical Transactions of the Royal Society of London B*, 292: 57-64.

Whybrow, P. J. & Bassiouni, M. A. (1986) The Arabian Miocene: rocks, fossils, primates and problems. In Else, J. G. & Lee, P. C., eds. *Primate Evolution*, pp. 85-91. Cambridge: Cambridge University Press.

The Evolution of Archaeological Perceptions of the First Americans: Historical Implications and Paradigms Lost

By Alvah M. Hicks

Archaeologically based skepticism surrounding "pre-Clovis" occupation of the Americas confounded the development of novel theories (including the paradigm of back-migration identified by Boas and his colleagues), long before the acceptance of even a Holocene human occupation in 1927.[1] The "Clovis First" paradigm continued the tradition of overshadowing hypotheses entailing "pre-Clovis" Ice Age habitation of the Americas *until* Monte Verde surfaced, gaining overwhelming acceptance in 1998 (Dillehay). These inaccurate archaeological perceptions have complicated (i) the dating and/or acceptance of "Amerind" (as a single linguistic group), as well as (ii) genetic identification in Northeast Asia of proposed "founding mtDNA lineages," and (iii) ideas of back-migration into Siberia by Athabaskans and/or Eskimos that could offer untested alternatives to traditional evaluations (Hicks 1998). Today scientific acceptance of mid-Pleistocene human occupation of the Americas beckons us to entertain ideas of migrations both into and out of the Americas.[2]

There has been a plethora of recent articles (some from scientific journals, some enclosed in popular publications like *Newsweek*) suggesting that Europeans were the "first" inhabitants of the Americas. This alternative, promoted by a few die-hard advocates of the "Clovis First" model, would have the first Americans coming from Europe across the Bering Strait before a later group, today's Amerindians, arrived and replaced them. I would like to identify inconsistencies with the – now archaic – "Clovis First" hypothesis by incorporating pre-Clovis Amerindians into the equation. Simply, there are viably significant geographic, linguistic, archaeological, and genetic correlates that would more accurately link a later, separate, post-Ice Age demic diffusion of Europeans with Upper Paleolithic technologies into earlier pre-Clovis Amerindian populations already inhabiting the Americas.[3]

The "Clovis First" paradigm began with the earliest evidence of an American Paleolithic stage, i.e., "Paleoindian Traditions."[4] Since "Fluted Points" are unique to the Americas, with no other evidence of refined Paleolithic tools predating them, archaeologists have long hypothesized that Eurasians equipped with Upper Paleolithic (UP) industries were the first Americans. Moreover, since these distinctive tool types first begin to appear at the end of the last Ice Age, an archaeological consensus had (or has, since there is still disagreement) categorically refuted any earlier evidence, inasmuch as all reputed pre-Clovis sites lack evidence of refined stone tools (Pearson 1997). Yet, today, we are finally beginning to accept that there were *indeed* Pleistocene Native American ancestors, leaving little, if any, evidence to link them with Old World Upper or Late Paleolithic people (Krieger 1964; Wormington 1957). Could UP-equipped people migrating from northeast Asia into North America, near the end of the last Ice Age, have influenced pre-Clovis Amerindians in the development of "Paleoindian Traditions"? If so, are there any genetic correlates to support demic diffusion by Eurasian people – equipped with large-game hunting technologies – into pre-existing Amerindian populations?

Since it is now archaeologically sanctioned to entertain hypotheses entailing earlier pre-Clovis populations living in the Americas during the Pleistocene, new formulas must emerge in order to integrate linguistic, cranial, genetic, archaeological, and geographic data (Rogers et al. 1985b; Rogers 1985a). For example, since European mtDNA lineages are found exclusively in *northern* North American Amerind speakers, but *not* in aboriginal groups further south (Central or South America), it could be theorized that Europeans were assimilated into these, the northern-most Amerindians, following post-Ice Age migration from northeast Asia about 12,000 years ago.

> Our analysis confirmed that haplogroup X is present in both modern Native Americans and European populations. For the Native Americans, this haplogroup encompasses ~25% of the Ojibwa, 15% of the Sioux, 11-13% of the Nu-Chah-Nulth, 7% of the Navajo, and 5% of the Yakima. Thus, with the exception of the Na-Dene-speaking Navajo, *the distribution of this haplogroup among the Native Americans appears to be restricted to northern Amerindian populations.* (Brown et al. 1998, p. 1857) [emphasis added]
>
> Recent European genetic admixture cannot explain the presence of haplogroup X in the Amerindians. *First, if the occurrence of haplogroup X were the result of female gene flow from Europeans, then other, more common European mtDNA haplogroups should also be present in the northern Native Americans, and they are not.* Second, the Native American and European mtDNAs are very different and are connected only through an ancient common ancestor. Hence, Native American and European haplogroup X mtDNAs diverged long ago. (*Ibid.,* p. 1857) [emphasis added]
>
> A coalescence time of 12,000-17,000 years ago could be interpreted as a rapid re-expansion, or, alternatively, *as an independent and late arrival of haplogroup X mtDNAs into the Americas.* (*Ibid.,* p. 1859) [emphasis added]

The genetic data would seem to indicate that European Type X mtDNAs represent a later, isolated migration. Moreover, pre-Clovis Amerindian populations must have been here before the assimilation of Europeans, since the European X-marker is found exactly where later-arriving Eurasians would have first encountered "pre-Clovis people" already inhabiting the Americas. Conversely, if the "first" Americans were Europeans, then one would expect haplogroup X mtDNAs to have survived in Central and South America, where they (haplogroup X mtDNAs) are implicitly absent. Europeans, accordingly, could not have been the *first* native people of the Americas. Rather, the identification of Eurasian mtDNA haplogroups X and J (Stone & Stoneking 1998) hints at a peaceful assimilation into, principally, the northernmost Amerindians, who genetically retain evidence of the point of contact between previously geographically isolated New and Old World Pleistocene populations.

The uniquely Amerindian development of fluted stone tools (Clovis and related "Paleoindian Traditions") could be directly attributed to this separate, independent, European migration, the result being post-Ice Age diffusion of Old World UP "know-how" into pre-existing pre-Clovis populations (Muller-Beck 1966). Since few, if any, archaeological components defining the Old World Paleolithic are found at Monte Verde (or any other pre-Clovis sites), it should be assumed that an outside (Old World) influence contributed to "Paleoindian" traditions, and the increased production of archaeological signatures accompanying the use of Paleolithic tools.[5]

We should remember that J.H. Greenberg (1987) was cautious when he dated "Amerind," and its 11 language stocks found in North and South America, basing the arrival of his first of three Native American language groups (Amerind) on a chronological generation that conformed to the *archaeological consensus* of the time (i.e., "Clovis first," being < 12,000 ybp).

> It [Amerind] may plausibly be connected with the Paleo-Indian (Clovis) culture, which dates back at least 11,000 to 12,000 years. ... Although we have presented linguistic criteria to establish a relative chronology of the three migrations, we have considered only archaeological correlations as a source for an absolute chronology. (Greenberg 1987, pp. 333-334)

The significance of a pre-Clovis occupation of the Americas *is* greater time-depth, compatible with the detection of extensive genetic and linguistic diversity in Native American populations. The reality of a greater time-depth must today incorporate the prospects of a post-Ice Age assimilation of once-isolated Eurasian populations into earlier pre-Clovis Amerindian tribal groups. New migration scenarios, compatible with pre-Clovis habitation of the Americas, must correlate geographic, genetic, linguistic, and archaeological data. One test could integrate both Eurasian expansion into, and Amerindian expansion back out of, the Americas (Boas 1905; 1910), after the last Ice Age. The recognition of a substantial pre-Clovis Amerindian population celebrates, as part of the equation, the peaceful assimilation of Eurasian people who may have carried with them to the Americas Upper Paleolithic industries following the end of the last Ice Age.

References:

Boas, F. 1905. The Jesup North Pacific Expedition. In *Proceedings of the Thirteenth International Congress of Americanists,* pp. 91-100. Easton, PA: Eschenback.

Boas, F. 1910. Ethnological Problems in Canada. *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 40: 529-539.

Brown, Michael D., S.H. Hosseini, A.Torroni, H. Bandelt, J.C. Allen, T.G. Schnurr, R. Scozzari, F. Cruciani, and D.C. Wallace. 1998. MtDNA Haplogroup X: An Ancient Link between Europe/Western Asia and North America? *American Journal of Human Genetics* 63: 1852-1861.

Greenberg, Joseph H. 1987. *Language in the Americas.* Stanford: Stanford University Press.

Hicks, Alvah M. 1998. Alternative Explanations for Similarities between Native Americans and Siberians. *Human Biology* 70.1: 137-139.

Hoffecker, J.F., W.R. Powers, and T. Goebel. 1993. The Colonization of Beringia and the Peopling of the New World. *Science* 259: 263-287.

Krieger, Alex D. 1964. Early Man in the New World. In *Prehistoric Man in the New World,* ed. by J.D. Jennings and E. Norbeck, pp. 23-84. Chicago: University of Chicago Press.

Krupnik, Igor. 1998. Jesup Genealogy: Intellectual Partnership and Russian-American Cooperation in Arctic/North Pacific Anthropology. Part 1. From the Jesup Expedition to the Cold War, 1897-1948. *Arctic Anthropology* 35.2: 199-226.

Muller-Beck, H. 1966. Paleohunters in America: Origins and Diffusion. *Science* 152: 1191-1210.

Pearson, Georges A. 1997. Further Thoughts on Clovis Old World Origins. *Current Research in the Pleistocene* 14: 74-76.

Rogers, Richard A. 1985a. Glacial Geography and Native North American Languages. *Quat. Res.* 23: 130-137.

Rogers, Richard A., L.A. Rogers, R.S. Hoffmann, and L.D. Martin. 1985b. Native American Biological Diversity and the Biogeographic Influence of Ice Age Refugia. *Journal of Biogeography* 18: 623-630.

Stone, Anne C., and Mark Stoneking. 1998. MtDNA of a Prehistoric Oneota Population: Implications for the Peopling of the New World. *American Journal of Human Genetics* 62: 1153-1170.

Wormington, H.M. 1957. *Ancient Man in North America.* (Denver Museum of Natural History, Popular Series No. 4, 4th Ed.) Denver.

Notes:

1. Inaccurate archaeological assessments championed by Aleš Hrdlička (that humans arrived in America less than 4,000 years ago: see Krupnik 1998) prematurely dismissed the Jesup Expedition's paradigm of back-migration into Siberia.
2. The diffusion of Upper Paleolithic technologies into pre-existing Amerindian populations, and the subsequent movement of Paleoindian traditions, with their distinctive fluted points, into deglaciated North America, suggests that migrations went in both directions following the end of the last Ice Age.
3. Do Nostratic and Cherokee represent a plausible linguistic correlation? (Ruth Holmes, personal conversation)
4. The term "Paleoindians" should not be used to define earlier "pre-Clovis" inhabitants, since the behaviors separating mid-Pleistocene Amerindian occupations lack evidence of advanced stone tools. Fluted points did not, in all probability, evolve from pre-Clovis lifeways. Rather, the sudden, independent, and coincidentally post-Ice Age development of Paleoindian traditions by Amerindians suggests contact with Eurasian Paleolithic people after the Eurasians migrated into unglaciated North America.
5. The sudden appearance of Upper Paleolithic-like lithic tools at the end of the last Ice Age is the crux of the "Clovis First" hypothesis, having defined for most archaeologists, though inaccurately, the initial phase of human occupation of the Americas.

Editorial:

In Honor of Jürgen Pinnow:Comparanda Dene-Caucasica

On the Occasion of his 75th Birthday
22 January 2000

Veteran Long Ranger Jürgen Pinnow lives on the Frisian island of Sylt. He recently sent me a Christmas letter, dated 17-12-99, thanking me for *Long Ranger* 32 (Fall 1999), and saying that he was retired from work on North American languages, but is now intensively studying the dying German dialects of the Danzig area, his home turf. For those who may be unfamiliar with Pinnow's work, he was first widely known as an authority on the Austroasiatic languages of India and Southeast Asia (see the references listed in *Mother Tongue* II, p. 150), and later he became one of the major authorities on the Na-Dene languages of North America (Haida, Tlingit, Eyak, and the Athabaskan family). Pinnow has remained one of the staunch supporters of the inclusion of Haida, clearly the most divergent member, in the Na-Dene family.

Pinnow is a Long Ranger. That is, he allows himself to think and hypothesize about distant relationships between the traditionally accepted language families. He thinks there is evidence for remote relationships between Na-Dene and certain other language families. However, his ideas (as expressed in Pinnow 1976 and 1990) do not precisely coincide with the Dene-Caucasian hypothesis (Na-Dene + Sino-Tibetan + Yeniseian + Caucasian + Burushaski + Basque), but are more similar to those of Morris Swadesh. Swadesh envisioned a vast linguistic network that connects all the languages of the world. Pinnow's tentative diagram (Pinnow 1990: 23) is of this type, in which Na-Dene is connected by solid lines with Sino-Tibetan to the west, and with (Greenberg's) Almosan-Keresiouan in North America (i.e., Wakashan, Salishan, Algic, Siouan, Iroquoian, etc.) to the east.

Some historical linguists accept only, or primarily, lexical evidence for the genetic affinity of languages, others accept only, or primarily, morphological (grammatical) evidence, though a combination of these types of evidence is the ideal. Pinnow's arguments are backed up by volumes of evidence, where, for example, every recorded word and sentence in the Haida language is painstakingly documented, analyzed, and compared with Tlingit, Eyak, and Athabaskan. (*Das Haida als Na-Dene-Sprache,* in *Abhandlungen der Völkerkundlichen Arbeitsgemeinschaft,* Heft 43, 44, 45, 46, 1985, Nortorf, Germany.) In my opinion, this work fully vindicates Edward Sapir's classification, with Haida as the "outlier," or most remote member, of the Na-Dene family. He has also (Pinnow 1976) documented the history of Na-Dene scholarship, from Rezanov and Buschmann, Boas and Sapir, to Hoijer and Krauss, though it now needs updating since 1975.

Probably less known is that Pinnow has supplied some of the most convincing evidence supporting the "Sino-Dene" hypothesis. This is the hypothesis linking Sino-Tibetan and Na-Dene, usually associated with Sapir (see Bengtson 1994). Following up on Sapir's declaration to Alfred Kroeber that Classical Tibetan "is startlingly Nadene-like," Pinnow has assembled some precise grammatical comparisons between Sino-Tibetan (usually Tibetan) and Na-Dene:

Sino-Dene Morphological Parallels

In his history of Na-Dene research (*Geschichte der Na-Dene Forschung,* 1976, pp. 94-105), Pinnow not only reports on the Sino-Dene studies of others (Sapir, Shafer, Swadesh), but also offers some interesting Sino-Dene morphological parallels of his own.

Regarding nouns, Pinnow outlines what appears to be an underlying Dene-Caucasian tendency to use possessive prefixes, as well as the concept of the unpossessed form (with prefix ***ʔa-**). Pinnow quotes the striking example cited by Paul K. Benedict, comparing Tibeto-Burman ***a-** with Navajo *ʔa-* (and Abkhaz *a-*):

> Tibeto-Burman ***a-myak** 'eye' (unpossessed) > Kachin *ə-myi* 'eye' (and Miri *ə-mik,* Lepcha *ă-mik*, etc.);
> Cf. Navajo *ʔa-nááʔ* '(someone's) eye' (unpossessed, as opposed to *ši-nááʔ* 'my eye', etc.);
> Cf. Abkhaz *a-bla* 'eye' (unpossessed, as opposed to *sə-bla* 'my eye', etc.)

The idea here is to compare the *first* element in each word: the unpossessed prefix *(ʔ)a-*. (On the words for 'eye' themselves: Tibeto-Burman ***myak** and Navajo *-nááʔ* < PAth ***(n-)we·ɢ-** = Tlingit *wàq* = *wàɢ* 'eye' are probably cognate, but the Caucasian word is of separate origin.) Pinnow expands on the subject, pointing out that there are other precise parallels between Sino-Tibetan and Na-Dene possessive prefixes, e.g:

Kachin **n**-*wa* 'thy father'	=	Navajo **ni**-*ǯéʔé*	'thy father'
Kachin **kə**-*wa* 'his father'	=	Navajo **ha**-*ǯéʔé*	'his father' (4th person)
Dimasa **bu**-*gur* 'skin' (< **bo-gur* his skin')	=	Navajo **bi**-*ǯéʔé*	'his father' (3rd person)
Meitei **mə**-*pa* 'his father'	=	Hupa, Sarsi **mi-**	(3rd person possessive)

Regarding pronouns, Pinnow points out a very precise first and second person parallel:

(Sino-Tibetan) Dhimal *ka* 'I' / *na* 'thou'
(Proto-Na-Dene) ***χa** 'I' / ***na** 'thou' (χ = unvoiced uvular fricative)
> Tlingit *χ(a)-* 'I' / Navajo *n(i)-* 'thou', etc.

The question here is whether Dhimal in fact preserves the Proto-Sino-Tibetan paradigm, which almost everywhere else reflects PST ***ŋā-** 'I' / ***nă-** 'thou' (according to Peiros and Starostin [1996]). Peiros and Starostin (ST V:134) explain Dhimal *ka* 'I' (and Tibetan *kho-bo* 'I, me', Lushai *ka* 'me, my', etc.) as secondary compounds of ***k(j)ă-** 'this, he' + ***-ā-** 'I'. In any case, the resemblance of second person PST ***nă-** 'thou' and PND ***na-** 'thou' is clear.

On the subject of verbs, we are virtually restricted to Classical Tibetan, the only well-known Sino-Tibetan language that preserves much of the (presumed) Dene-Caucasian verbal morphology. When Sapir declared that Tibetan was "startlingly

Nadene-like,” he was thinking of several typological parallels. One is the presence in both languages of verbal stem variants (allomorphs) distinguished by ablaut, for example:

	I (imperfect, inchoative)	P (perfect)	F (future)	O (optative)	
Tibetan	*sel*	*b-sal-d*	*b-sal*	*sol-d*	‘to cleanse, remove’
Tlingit	*hén*	*hùù*	*húú*		‘to swim’
	t’úk^{w}	*t’ùùk*^{w}	*t’úúk*^{w}		‘to shoot (an arrow)’
Chipewyan	*-t’ù*	*-t’òγ*	*-t’ù*		‘to sting’
Navajo	*-t’òh*	*-t’óóh*	*-ɬ-t’ò*		‘to shoot (an arrow)’

Both languages also feature at least two prefix positions before the verb stem, for example, in Tibetan: *s-gul* ‘to move, set in motion’, (perfect form) *b-s-gul:*

b-	*-s-*	*gul*
tense prefix (perfect)	transitive prefix	verb root 'to move'

In most Na-Dene languages, there are several prefix positions, e.g. in Tlingit:

ʔad-wuu-si-tììn 'he saw something'

ʔad-	*-wuu-*	*-si-*	*-tììn*
pronoun object 'something'	tense prefix (perfect)	'classifier' transitive	verb root 'to see'

So far we have only mentioned *typological* similarities, which are, of course irrelevant for the genetic classification of languages. Indo-European and Semitic languages also have ablaut, for example. It is only resemblances of sound *and* meaning that are useful for genetic classification. Pinnow therefore goes on to show that Sino-Tibetan and Na-Dene have cognate morphological elements as well as similarities in the sequence of those elements. In the examples just mentioned, we see the same sequence of morphological elements, as well as probable cognacy of the elements themselves:

Tibetan	*b-*	*-s-*	*-gul*
	tense +	trans. +	verb root
Tlingit (*ʔad*)	*-wuu-*	*-si-*	*-tììn*

Pinnow proposes that the Tibetan perfect prefix *b-* may be cognate with the Tlingit perfect prefix *-wu(u)-*. The sound correspondence is plausible, since Tlingit has no phoneme /b/, and in fact we find that all Dene-Caucasian labials tend to merge into one phoneme (usually reconstructed as ***w)** in Na-Dene, e.g.:

> PST ***mjVk** 'eye' (Moshang *mak*) = Tlingit *wàq (wàɢ)* 'eye'
> PST ***miał** 'sleep' = PA ***wəł** > Navajo *bił,* Hupa *mił* 'sleep'
> PST ***pāk,** Old Chinese***pāk** 'wide, long, ample' = Tlingit *wùq'* 'wide, broad'
> PST ***br[ē]ł** 'to divide, separate' = Tlingit *wál', wààl', wáál* 'to break'

Pinnow also suggests that the following elements in the Tibetan-Tlingit comparison, above, namely the transitive affixes (Tibetan **s-** = Tlingit –**s**(i)-) are cognate. The same morpheme is found in Haida, as in this comparison cited by Sapir:

> Haida *s-k'ul, s-k'al* 'to boil, roast' (vs. unprefixed form in Athabaskan: Sarsi *-kạl* 'water boils'):
> cf. PST ***kōł** 'to boil' > Tibetan I *s-kol,* PF *b-skol* 'to boil' (trans.) v. I *ã-khol,* P *khol* 'to be boiled';

Incidentally, Burushaski also has this –**s**- transitivizer:

> Burushaski *γól-/γul-* 'to burn' (intransitive) vs. –**s**-*qul-* 'to burn' (transitive)

In the case of Haida **s**-*k'ul* ~ Tibetan **s**-*kol* ~ Burushaski –**s**-*qul-,* both prefix and root are very probably cognate. Cf. also:

> Burushaski *du-γáṭ* 'chosen, selected, separated out' vs. *d-s-qaṭ* 'to choose, select' ~ Haida *s-q'at* 'to recognize', *s-q'at'-áa* 'to learn' < **q'ad* 'to know' (to recognize is to select); one is tempted to add Tibetan *skad (s-kad)* 'voice, sound, speech; to say, tell'

Quoting Benedict again, Pinnow compares Tibeto-Burman ***m-** (middle voice/ durative, intransitive, reflexive) with the Athabaskan morpheme manifested as Navajo and Mattole **bi**-, Hupa and Sarsi **mi-:**

> e.g., Navajo **bi** – *di* – *ɫ* – *teeh* 'he begins to carry him'
> 'him' 'begin' (classifier) 'carry'

From the Tibetan side, Pinnow compares forms such as **m**-*nal-ba* 'to sleep', **m-***thoŋ-ba* 'seen'. (Compare also the third-person possessives cited above.)

Comparative Dene-Caucasian morphology is a potentially rich field for future research, particularly when we bring evidence from the other languages, such as Burushaski, Yeniseian, Caucasian, and Basque. We need to make good use of the meticulous work Jürgen Pinnow has begun.

For Jürgen Pinnow:

Some Na-Dene – Dene-Caucasian comparisons:

John D. Bengtson January 20, 2000

1. Na-Dene: PND ***χUʔ** ‘tooth’ > Tlingit *ʔùuχ,* Eyak *χuu-ł,* PPA ***χwu(u)ʔ** > Chipewyan *xù-, -γù-, -γú-,* Mattole *-γwoʔ,* Navajo *-γooʔ*, etc. ‘tooth’

~ Yeniseian: PY ***χowe** ‘mouth’ > Ket *qō,* Yug *xo,* Kott *hōpi*, etc. ‘mouth’ (SSEJ 302)

~ Sino-Tibetan: PST ***khʷə(H)** ‘mouth’ > Old Chinese ***khōʔ**, Dimasa *khu,* etc. (ST V:107)

~ Caucasian: cf. PNC ***χwɨm(V)ṗV** ‘gulp, mouthful’ > Khinalug *χob,* etc. (NCED 1082; old compound of ***χwV(ʔ)** + ***m(V)ṗV** ?)

~ Basque ***a-x(ʷ)o** ~ ***a-γ(ʷ)o** > *aho* (BN,L,Z) ~ *ao* (AN,G,R,B) ~ *ago* [aγo] (AN,G,R,B) ~ *abo* [aƀo] (B) ‘mouth’

2. Na-Dene: PA ***łu:t’** > Hupa *łoh, -lood-eʔ* ’scab’, Mattole *–lood-eʔ* ‘scar’, Chipewyan *łùr* ‘scab, chunks of ice’, Carrier *łud* ‘crust on wound’, Kutchin *łíd* ‘scar’, Navajo *łóód, -lóód* ‘sore’

~ Sino-Tibetan: PST ***lVt** ‘mucus, phlegm’ > Old Chinese ***slits** ‘drivel from the nose’, Tibetan *lud* ‘phlegm, mucus; manure, dung’, Jingpo *səlat*^{2} ‘sweat’ (ST III:51)

~ Caucasian: PEC ***ł̱wirdɨ** ‘manure; pus’ > Avar *xʷerd* ‘pus, matter’, Bezhta *łär* ‘sheep’s dung’, Agul *furd* ‘manure, dung’, etc. (NCED 763)

~ Basque *lirdi* (G) ~ *lerde* (AN,BN,G) ‘drivel, saliva’

§ Semantics: originally ‘secretion/excretion’; cf. no. 14: PA ***ǯeχʔ**, etc., below.

3. Na-Dene: Tlingit *ʒas* ‘skin’, *ʒáás* ‘thong’, Eyak ku-*sic’* ‘fish-skin’, Galice *saas* ‘skin (of animals)’, Hupa *–sic’* ‘skin, bark’, etc.

~ Yeniseian: PY ***sā̄s** ‘(animal) skin’ > Ket *śaś* ~ *śäśi* ~ *śa:ś,* Yug *sa:ʰs,* Kott *šēt,* etc. (SSEJ 270)

~ Caucasian: PCircassian ***č́’əśʷa** ‘human skin, body’ > Adyge *č́’əśʷa,* Kabardian *ś’əfa*

4. Na-Dene: Haida *χay* ‘(blood) vein’; PA ***χa:y** ‘root’ > Chipewyan *xai, -γay-* ‘root’, Hupa *xay* ‘root (of conifer)’, etc.

~ Caucasian: PNC ***x̱wɨʔrV** ‘vein’ > Chechen *pχa* ‘vein’, Lak *x:ʷa* ‘sinew, tendon; string’, etc. (NCED 1064)

~ Basque *erro* ~ *herro* ‘root (also, of teeth); teats (of udder); ray (of sun)’

5. Na-Dene: Haida *s-qál* ‘shoulder’, Eyak *ɢəlaʔ* ‘shoulder, shoulder-blade’
~ Yeniseian: PY ***χɔlab** ‘side, half’ > Ket *qɔl^{y}ap^{5}*, Kott *halap* (SSEJ 304)
~ Sino-Tibetan: PST ***qālH** ‘back, small of the back’ > Tibetan *s-gal*-pa ‘small of the back’, Meyol *kal*-po ‘shoulder’, Lushai *ēl* ‘small of the back, etc. (ST V:152)
~ Caucasian: PEC ***q̱ĭłʔi** ‘elbow, arm, wing’ > Lak *qa* ‘wing’, Archi *χol* ‘hand’, etc. (NCED 895); and/or PNC ***ɢw[ü]łŏ** ‘side’ > Lezgi *q:^{w}al,* etc. (NCED 472)
~ Burushaski *galgí* ‘wing; fletching of arrow; side part of saddle’
~ Basque *hegal* ‘wing’

6. Na-Dene: Tlingit *č’íχ’ɪ* ‘index finger’, Eyak *cəq’-s* ‘fingers, toes’
~ Yeniseian: PY ***təʔq** ‘finger’ > Ket *tʌʔq,* Kott *thok, thox(g),* Pumpokol *tok,* etc. (SSEJ 283)
~ Sino-Tibetan: Classical Tibetan *m-ʒug-gu = m-dzug-gu,* Burig *zūh* ‘fingers, toes’
~ Caucasian: ? Andi *c̄’eka* ‘finger’
~ Basque (B) *zuku+atz* ‘index finger’
§ Sporadically attested!

7. Na-Dene: Eyak *ǯiʔƛ̆’g* ‘navel’
~ Yeniseian: PY ***tɨl-** ‘navel’ > Ket *tɨɨl^{y}*, Yug *tɨl* (SSEJ 286)
~ Sino-Tibetan: PST ***ćial** ‘navel’ > Old Chinese ***ʒəj** ‘navel’, Lushai *thial* ‘gizzard’ (ST IV:49)
~ Caucasian: Tabasaran *c’ul* ‘navel’, Dargwa *č’ilič’a* ‘navel’ (< ***ǯili-k’V,** cf. Eyak, above, and Basque [čilko])
~ Basque (c) *zil*-bor ‘navel’ ~ (B) *zil* ~ (AN,G) *txilko* [čilko] ~ (AN,BN,L) *xilko* [šilko] ‘navel’

§ Several apparent phonological irregularities (especially in Caucasian), mainly caused by diminutive suffixes, assimilation, and affective palatalization. The first of three Dene-Caucasian ‘navel’ etymologies; see nos. 8 and 9, below, and Bengtson 1998.

8. Na-Dene: Eyak *c’aaʔ*‘umbilical cord’, PA ***c’e:ʔ** ‘navel’, diminutive ***c’e:q’** > Tututni *c’eʔ*, Hupa *–c’eeq’,* Sarsi *–c’ák’-*, Navajo *–c’ééʔ*, etc. ‘navel’
~ Burushaski (Y) *–sú,* pl. *–súmu ~ -súmo,* (H) *–súi,* pl. *súimuc* ‘umbilical cord, navel’
~ Caucasian: PEC ***ʒŏnʔŭ** ‘navel’ > Khinalug *c’um,* Tindi *c:ũũ*, Dargwa *zu,* PNakh ***c’an-k’u** (cf. Sarsi, above) > Chechen *c’onga*, etc. ‘navel’ (NCED 1096)

9. Na-Dene: Haida *s-gíl* 'navel, umbilical cord', Tlingit *kùl* 'navel, umbilical'

~ Sino-Tibetan: Tibetan *d-kyil-ma* 'middle, center', Ladwags *s-kyil,* Balti *s-kil* 'center'

~ Burushaski (Yasin) *–skil* ~ *-skul* ~ (Hunza) *–skil* ~ *-ski* 'Gesicht; Oberfläche; Seite (eines Kastens usw.)' ['face; surface; side (of a chest, etc.)']; *–skil -ce* 'before' (adv.)

§ Semantic change possibly 'navel > middle > half > side'

10. Na-Dene: Haida *s-tl'a*-wúl 'rectum', *s-tl'áa*-nǯaaw 'toilet paper'; Navajo *ƛ'ááh* 'bottom', *-ƛ'aaʔ* 'buttocks', Hupa *-ƛ'aʔ* 'buttocks, bottom', Carrier *–tɬ'ah* 'posterior'

~ Burushaski (Hunza) *tano* 'lower end of an animal's gut (= colon)', *tan* ~ *tano* 'bastard (= base born)' < ***-ltan(o)**

~ Caucasian: PNC ***Hƛ̆'ŏnŭ** 'bottom' > Bezhta *õƛ'o,* Avar *t'ínu,* Lezgi *k'an*, etc. (NCED 590)

§ One of several etymologies involving lateral affricates. Cf. also nos. 12, 15.

11. Na-Dene: Tlingit *s'í* ~ *s'ɛ́* 'eyebrows', Eyak *c'ããχ* 'eyebrow', Chipewyan *–t'θ̃ì* (in *sɛ-t'θ̃ì-daγá* 'my eyebrows'), all < PND ***c'VN-**

~ Yeniseian: PY ***cəŋe** 'hair' > Ket *təŋə*, Yug *čəŋ*, Kott *heŋai,* etc. (SSEJ 213)

~ Sino-Tibetan: PST ***chām** 'hair (of head)' > Old Chinese ***s̱rām** 'hair', Tibetan ʔag-*chom* 'beard', Magari *čham* 'hair, wool', Kanauri *tsam* = *cam* 'wool', mik-*tsam* = mik-*cam* 'eyebrow', Garo mik-*sam* 'eyebrow', etc. (ST IV:19)

~ Caucasian: PEC ***c'ɦwäme** 'eyebrow' > Chechen *c'oc'q'am,* Kryts ʕül-*c'äm* 'eyebrow', etc. (NCED 364)

~ Basque *zamar* (***sama-r**) (Z) 'longs poils d'une bête de somme mal soignée' ['long, unkempt hair of a pack-animal'] ~ 'lock of wool', etc. (cf. Magari, Kanauri, above)

§ The specialized meaning 'eyebrow(s)' became exclusive in Na-Dene andCaucasian, but only partially in Sino-Tibetan (Kanauri, Garo).

12. Na-Dene: [***ƛ̆'aχ** = ***t̯ɬ̆'aχ**] Navajo *ƛ'ah* 'temples, sides of the face, side of head', Carrier *–tɬ'a-* (in *–tɬ'a-* 'temples', *–tɬ'a*-yus 'side whiskers')

~ Sino-Tibetan: PST ***[ƛ̆]ĕk**^{w} 'back' > Tibetan *ltag* 'the back part of the neck, nape; the back', Lushai *thluk* 'the back of the head, occiput', etc. (ST III:62)

~ Caucasian: PEC ***ƛ̆'arq'**^{w}**ĕ** 'forehead' > Khwarshi *ƛ'oq'o* 'forehead', Andi *laq'*^{w}*ara* 'cap', etc. (NCED 775)

~ Basque (B,G) *loki* 'temple (of the head)' ~ (R) *lokun* ~ (AN,BN,G,L,Z) *lo* 'temple'

§ Original 'side of the head', specialized to 'temples' in Na-Dene and Basque, 'back of head' in Sino-Tibetan, 'front' = 'forehead' in Caucasian.

13. Na-Dene: Haida *skuts = skuc (s-kuc)* ‘bone’
~ Sino-Tibetan: PST ***kūt** > Old Chinese ***kūt** ‘bone’ (ST V:75 – ***kūt** ‘hand’ may be a homonym)
~ Caucasian: PEC ***kŏc’a** ‘a kind of bone’ > Hunzib *k’oc’u* ‘back of the head’, Agul *k’ac’* ‘vertebra’, etc. (NCED 698)
~ Basque ***khoč**, as in (Z) gar-*khotx* [garkhoč] ‘nape’

14. Na-Dene: PA ***ǯeχʔ** ‘pitch, gum’ > Navajo *ǯééh,* Hupa *ǯeh, ǯeewʔ,* Chipewyan *dzέ, -dzὲγέ*, etc.
~ Yeniseian: PY ***ǯik** ‘resin, pitch, tar’ > Ket *dik,* Yug *dʸik,* Kott *čīk,* etc. (SSEJ 310)
~ Burushaski (Hunza) *číki* ‘musk (of a person or plant)’, *číki*–muǯúr ‘musk weeping willow’
~ Caucasian: PNC ***č̱’ɨq’wA** ‘dung, ordure, dirt’ > Budukh *č’uq’* ‘eye secretion, earwax, mould, fish eggs’ (!), etc. (NCED 387)
~ Basque *zikin* ‘dirty’; (Z) *zikhina* [sikhĩ́ñã] ‘crasse que forme sur la peau la sueur desséchée’

§ For semantics: original ‘secretion/excretion (of person, animal, or plant)’. Cf. no. 2: PA ***łu:t’**, etc., above.

15. Na-Dene: Tlingit *ƛ’έχ* ‘brown fungus’, Eyak *ƛ’ɪχ* ‘grass, green’, PA ***ƛ’uχ** ‘grass’ > Navajo *ƛ’oh,* Sarsi *-ƛ’o-,* Galice *t’loh,* etc.
~ Sino-Tibetan: PST ***lŭH** ‘weed, to weed’ > Old Chinese ***luʔ** ‘weeds’, Lushai *thlo (thlawʔ)* ‘to weed, cultivate’, etc. (ST III:43)
~ Caucasian: PEC ***ʔwēƛ’V** ‘a kind of grass’ > Chechen *jol* ‘hay’, Tsakhur *ok’* ‘grass’, etc. (NCED 230)
~ Basque (c) *arda-* ~ (L) *hardo* ‘tinder (made from a kind of dried fungus)’ (cf. Tlingit, above)

§ All developments of metathetic variants such as ***ƛ’eHʷV** ~ ***Hʷeƛ’V** ~ ***Heƛ’ʷV.**

16. Na-Dene: Haida *ɢayt ~ ɢayt’áa* ‘ashes’; Tlingit *kέl’t’* ‘(wood) ashes’
~ Caucasian: PEC ***q̱’idV** ‘dust, soot’ > Akhwakh *q’:ẽt’a* ‘soot’, etc. (NCED 927)
~ Basque: *kedar* ‘soot’ ~ *k(h)eder* (BN,L) ~ *k(h)elder* (BN,L,Z), etc. ‘soot’

17. Na-Dene: Haida *sing = sɪŋ* ‘sky, day’; PA ***səŋ**$_2$**’** ‘star’ > Ahtna *son’ = sonʔ*, Kaska *sʊn,* Chipewyan *θə́n,* Carrier *səm,* Navajo *sǫ̀’ = sǫ̀ʔ*, etc. ‘star’
~ Sino-Tibetan: PST ***sēŋ** ‘star’ > Old Chinese ***sēŋ** , Hruso li-*tsoŋ* = li-*coŋ,*

Tengsa lŭ-*tiŋ tiŋ*, PKiranti ***saŋ**, etc. ‘star’ (ST IV:99)
~ Burushaski (Yasin) *asúmun* ~ *asúmen* ‘star’ ~ (Hunza) *así* ‘star’, *asímuc* ‘stars’

18. Na-Dene: Haida *su* ‘to say’, Tlingit *–sa, -sá, -sén* ‘to name’ (‘to breathe’ is probably a homonym), *sa* ‘name, voice’
~ Burushaski *sén-, sí-* ‘to say, call, name’, *sén*-as ‘named’
~ Basque (c) *izen* (i-*zen*) ~ (B) *uzen* (u-*zen*) ‘name’

* * * * * *

Abbreviations:

AN	Alto Navarro (Basque dialect)
B	Bizkaian (Basque dialect)
BN	Basse Navarre = Behe-Nafarroa (Basque dialect)
c	Common (Standard) Basque
G	Gipuzkoan (Basque dialect)
H	Hunza (Burushaski dialect)
L	Lapurdi = Labourdin (Basque dialect)
NCED	*North Caucasian Etymological Dictionary* (see Nikolayev & Starostin, 1994)
P	Proto-
PA	Proto-Athabaskan
PEC	Proto-East Caucasian
PNC	Proto-(North) Caucasian
PND	Proto-Na-Dene
PPA	Proto-Pre-Athabaskan
PST	Proto-Sino-Tibetan
PY	Proto-Yeniseian
R	Roncalés (Basque dialect)
SSEJ	*Sravnitel’nyj slovar’ enisejskix jazykov* (see Starostin, 1995)
ST	Sino-Tibetan (see Peiros & Starostin, 1996)
Y	Yasin = Werchikwar (Burushaski dialect)
Z	Zuberoan = Souletin (Basque dialect)

REFERENCES

Bengtson, J.D. 1994. "Edward Sapir and the 'Sino-Dene' Hypothesis." *Anthropological Science* 102.3: 207-230.

Bengtson, J.D. 1998. "Dene-Caucasian 'Navel': Some Proposed Etymologies." *Dhumbadji!* 4.1: 86-90.

Blažek, V. and J.D. Bengtson. 1995. "Lexica Dene-Caucasica." *Central Asiatic Journal* 39.1: 11-50; 39.2: 161-164.

Kuipers, A.H. 1975. *A Dictionary of Proto-Circassian Roots.* Lisse, Netherlands.

Nikolayev, S.L. and S.A. Starostin. 1994. *A North Caucasian Etymological Dictionary.* Moscow: Asterisk Press.

Peiros, I. and S.A. Starostin. 1996. *A Comparative Vocabulary of Five Sino-Tibetan Languages.* Melbourne: University of Melbourne.

Pinnow, H-J. 1964. "On the Historical Position of Tlingit." *International Journal of American Linguistics* 30.2: 155-164.

Pinnow, H-J. 1966. *Grundzüge einer historischen Lautlehre des Tlingit.* Wiesbaden: Harrassowitz.

Pinnow, H-J. 1976. *Geschichte der Na-Dene Forschung.* Berlin: Indiana.

Pinnow, H-J. 1985. *Das Haida als Na-Dene-Sprache.* (Materialen zu den Wortfeldern und zu Komparation des Verbs = AVA, vols. 43-46.) Nortorf, Germany.

Pinnow, H-J. 1990. *Die Na-Dene Sprachen im Lichte der Greenberg-Klassifikation.* (= AVA, vol 64). Nortorf, Germany.

Starostin, S.A. 1995. "Sravnitel'nyj slovar' enisejskix jazykov." In *Ketskij sbornik (Studia Ketica),* ed. by S.A. Starostin. Moscow: Shkola Jazyki Russkoj Kul'tury.